

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 10:10:21 ; Search time 27 Seconds
(without alignments)
2231.389 Million cell updates/sec

Title: US-10-089-678-1

Perfect score: 6044
Sequence: 1 MSPNNQVEYILDASSSTSV.....MSETEGVNIESVELVLEE 1167

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/prodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/prodata/2/iaa/6C_COMB.pep.*
- 6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4007	66.3	1157	1	US-07-876-280-30
2	4007	66.3	1157	1	US-07-812-180A-2
3	4007	66.3	1157	1	US-08-315-468-2
4	4007	66.3	1157	3	US-07-941-650A-2
5	3662.5	60.6	1149	1	US-07-915-203-2
6	3662.5	60.6	1149	1	US-08-272-887-2
7	3662.5	60.6	1149	2	US-08-789-449-2
8	3152	52.2	1169	1	US-08-315-468-4
9	3123	51.7	1157	2	US-08-532-547-5
10	3123	51.7	1157	2	US-08-379-656B-5
11	3123	51.7	1157	3	US-08-455-838-5
12	3123	51.7	1157	3	US-09-019-809-5
13	3123	51.7	1157	4	US-09-471-177-5
14	3100.5	51.3	1156	4	US-09-002-285-72
15	3100.5	51.3	1156	4	US-09-589-477-72
16	3100.5	51.3	1156	4	US-09-661-322A-28
17	2825.5	46.7	1150	4	US-09-002-285-74
18	2825.5	46.7	1150	4	US-09-589-477-74
19	2715	44.9	1169	1	US-08-542-921-2
20	2715	44.9	1169	2	US-08-880-685-2
21	2715	44.9	1169	2	US-08-880-684-2
22	2710	44.8	1134	4	US-09-002-285-76
23	2710	44.8	1134	4	US-09-589-477-76
24	2549	42.2	1229	1	US-08-100-709-4
25	2549	42.2	1229	1	US-08-176-865-4
26	2549	42.2	1229	1	US-08-474-038-4
27	2549	42.2	1229	2	US-08-779-046-4

28 2549 42.2 1229 2 US-08-881-340-4 Sequence 4, Appli
29 2525 41.8 1227 3 US-09-053-549-2 Sequence 2, Appli
30 2519.5 41.7 1207 1 US-07-951-715A-7 Sequence 7, Appli
31 2519.5 41.7 1207 2 US-08-459-448A-7 Sequence 7, Appli
32 2519.5 41.7 1207 3 US-08-459-595A-7 Sequence 7, Appli
33 2519.5 41.7 1207 3 US-08-459-504B-7 Sequence 7, Appli
34 2519.5 41.7 1207 3 US-08-459-444-7 Sequence 7, Appli
35 2519.5 41.7 1207 3 US-09-053-549-8 Sequence 8, Appli
36 2519.5 41.7 1207 4 US-09-547-422-7 Sequence 7, Appli
37 2459 40.7 1227 1 US-08-448-170-8 Sequence 8, Appli
38 2459 40.7 1227 3 US-08-961-803-9 Sequence 9, Appli
39 2458.5 40.7 1138 1 US-07-973-320-2 Sequence 2, Appli
40 2457 40.7 1227 4 US-09-661-322A-63 Sequence 63, Appli
41 2454.5 40.6 1186 3 US-09-178-252-23 Sequence 23, Appli
42 2454.5 40.6 1186 4 US-09-828-660-23 Sequence 23, Appli
43 2452.5 40.6 1228 4 US-09-661-322A-38 Sequence 38, Appli
44 2441.5 40.4 1138 1 US-07-973-320-4 Sequence 4, Appli
45 2237.5 37.0 1177 3 US-08-754-490-14 Sequence 14, Appli

ALIGNMENTS

RESULT 1
US-07-876-280-30
; Sequence 30, Application US/07876280
; Patent No. 5262158
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M.
; APPLICANT: Cannon, Raymond J.C.
; APPLICANT: Bagley, Angela L.
; TITLE OF INVENTION: No. 5262158el Bacillus thuringiensis Isolates for
; TITLE OF INVENTION: Controlling Acarides
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/07/876,280
; FILING DATE: 19920430
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M/S 104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1157 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: kumamotoensis
; INDIVIDUAL ISOLATE: PS50C
; IMMEDIATE SOURCE:
; CLONE: E. coli NM522 (pMYC2320) NRRL B-18769
US-07-876-280-30

Query Match 66.3%; Score 4007; DB 1; Length 1157;
Best Local Similarity 66.1%; Pred. No. 2.3e-285;
Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;

QY 1 MSPNNQNEVEILDASSSTSVSDNSVRYPLANDOTTTLQNNYKDYLRMSEGENPELFGNP 60
DB 1 MSPNNQNEVEILDASSSTSVSDNSVRYPLANDOTTTLQNNYKDYLRMSEGENPELFGNP 60

QY 61 ETPFISSTVGTGIGVQVIGALGVFPAGQIASFYFIVGQLWPSSTSVWEMIMKQVED 120
DB 61 ETPFISSTVGTGIGVQVIGALGVFPAGQIASFYFIVGQLWPSSTSVWEMIMKQVED 120

QY 121 LIDOKITDVRKTAGLQGLGGLDVIYQKSLQWLENRNDTRARSVVQYIYALBDFV 180
DB 121 LIDOKITDVRKTAGLQGLGGLDVIYQKSLQWLENRNDTRARSVVQYIYALBDFV 180

QY 181 AKIPSPAISQOEYPLSVYAAQANLHLLLRDASIFGAEWGFTGPGISFYDROVTRTAQ 240
DB 181 AKIPSPAISQOEYPLSVYAAQANLHLLLRDASIFGAEWGFTGPGISFYDROVTRTAQ 240

QY 241 YSDYCVKWNVTGLDKLGTNAASWLYKHQFRREMTLLVLDLVALFPNYDTRTYPIETTAQ 300
DB 241 YSDYCVKWNVTGLDKLGTNAASWLYKHQFRREMTLLVLDLVALFPNYDTRTYPIETTAQ 300

QY 301 LTRVVDTPVFNRETSGGCRWSLNSDISFSEVESAVIRSHPLFDLISEHFYTRAG 360
DB 301 LTRVVDTPVFNRETSGGCRWSLNSDISFSEVESAVIRSHPLFDLISEHFYTRAG 360

QY 361 LPLNNTLEYLYWYGHISKYKNT-NASSALERNYGTITTSNKIKYDILANKDIFOVRSIGAD 419
DB 361 LPLNNTLEYLYWYGHISKYKNT-NASSALERNYGTITTSNKIKYDILANKDIFOVRSIGAD 419

QY 420 LANYAQVYGPVYASFLLDKNTGSGVGGFTYKSPHTTMQVCTQNYNTIDEIPPENE-P 478
DB 420 LANYAQVYGPVYASFLLDKNTGSGVGGFTYKSPHTTMQVCTQNYNTIDEIPPENE-P 478

QY 479 LSGYSHRLSHITSYFSGKNASSPARYGNLPVFAWTHRSADVNTVYSKTIQIPVVKAH 538
DB 479 LSGYSHRLSHITSYFSGKNASSPARYGNLPVFAWTHRSADVNTVYSKTIQIPVVKAH 538

QY 539 TLVSGTIVKPGFTGNILKRTSSGPLAYTSVKSPLSQRVRAIRVASTNNLFLVT 598
DB 539 TLVSGTIVKPGFTGNILKRTSSGPLAYTSVKSPLSQRVRAIRVASTNNLFLVT 598

QY 599 ISGTIYSINVNKTMKGGDLDLNTFPLATIGTAFPTFSNYSLSLTVGADSPAGSGVYVD 658
DB 599 ISGTIYSINVNKTMKGGDLDLNTFPLATIGTAFPTFSNYSLSLTVGADSPAGSGVYVD 658

QY 659 KPELIPVNATFEAEEDLVAKKAVNGLFTSKDALQTSVTDYQVNOANLVECLSDLYP 718
DB 659 KPELIPVNATFEAEEDLVAKKAVNGLFTSKDALQTSVTDYQVNOANLVECLSDLYP 718

QY 719 NEKRLMDVAKEAKRLVQARNLLQDTGFNRINGENGWGTSGTGLVEAGDVLFKDRSLRT 778
DB 719 NEKRLMDVAKEAKRLVQARNLLQDTGFNRINGENGWGTSGTGLVEAGDVLFKDRSLRT 778

QY 779 SAREIDPTPTLYQQIDSLKPYTRYKLGFGSSQDLKILIRHRANOIVKXVPDN 838
DB 779 SAREIDPTPTLYQQIDSLKPYTRYKLGFGSSQDLKILIRHRANOIVKXVPDN 838

QY 839 LLPDLVPSNCGGIDRCSEQQYVDANALENNGENGMSSDASHAFPHIDTGTIDNENT 898
DB 839 LLPDLVPSNCGGIDRCSEQQYVDANALENNGENGMSSDASHAFPHIDTGTIDNENT 898

QY 899 GIWVFKIPTNGYATLGNLELVEEGLPSGETLERAQQOQOQMDKMKRKGASEKAYVA 958
DB 899 GIWVFKIPTNGYATLGNLELVEEGLPSGETLERAQQOQOQMDKMKRKGASEKAYVA 958

QY 959 AKQADRLFDYQDQKLSNGVMSDMLAAQNLVQSIPYVYNDALPEIPGNMYTSFTLTN 1018
DB 959 AKQADRLFDYQDQKLSNGVMSDMLAAQNLVQSIPYVYNDALPEIPGNMYTSFTLTN 1018

QY 1019 RIQQANWLYDLRNAIPNGDFRNLGSLDMNATSDVNVQQLSDTSVLVIPNNNSOVSOQFTVQ 1078
DB 1019 RIQQANWLYDLRNAIPNGDFRNLGSLDMNATSDVNVQQLSDTSVLVIPNNNSOVSOQFTVQ 1078

QY 1079 PNYRYVLRVARTARKEGVGYVIRRGANQOTELTNICDDDTGVLSDADOTSITTKTVEFT 1138
DB 1079 PNYRYVLRVARTARKEGVGYVIRRGANQOTELTNICDDDTGVLSDADOTSITTKTVEFT 1138

QY 1139 PSTEQWIDMSETEGVFNIYESVELVLEE 1167
DB 1139 PSTEQWIDMSETEGVFNIYESVELVLEE 1167

RESULT 2

US-07-812-180A-2
; Sequence 2, Application US/07812180A
; Patent No. 5366892
; GENERAL INFORMATION:
; APPLICANT: Poncerrada, Luis R
; APPLICANT: Payne, Jewel M
; APPLICANT: Sick, August J
; TITLE OF INVENTION: No. 5366892el Coleopteran-Active Bacillus
; TITLE OF INVENTION: thuringiensis Isolate and a No. 5366892el Gene Encoding a
; TITLE OF INVENTION: Coleopteran-Active Toxin
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Roman Saliwanchik
; STREET: 2421 N.W. 41st Street, Ste A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/812,180A
; FILING DATE: 19920102
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, Roman
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1157 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: kumamotoensis
; INDIVIDUAL ISOLATE: PS50C
; IMMEDIATE SOURCE:
; LIBRARY: Lambdagem (TM) - 11 LIBRARY OF LUIS
; CLONE: 50C
; US-07-812-180A-2

Query Match 66.3%; Score 4007; DB 1; Length 1157;

Best Local Similarity 66.1%; Pred. No. 2.3e-285;

Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;

QY 1 MSPNNQNEVEILDASSSTSVSDNSVRYPLANDOTTTLQNNYKDYLRMSEGENPELFGNP 60

DB 1 MSPNNQNEVEILDASSSTSVSDNSVRYPLANDOTTTLQNNYKDYLRMSEGENPELFGNP 60


```

Db 121 LVQKLEKVKOKALAEKGLNALDVYQOOSLEDMLENRNDARTSRVSNQFIADLNPFV 180
Qy 181 AKIPSPAISSQEQEPLLSVYAQAANLHLLLRDASIFGAEWGFTPGBEISFYDQRTTAAQ 240
Db 181 SSIPSPAVSGHEVLLAVYAQAANLHLLLRDASIFGEEMGFTPGBEISFYDQRTTAAQ 240
Qy 241 YSDYCVKWTNGLDKLKGNTAASLWLYKHQFREMTELLVLDLVALPNNYDTRYPYETTAQ 300
Db 241 YSDYCVKWTNGLDKLKGNTAASLWLYKHQFREMTELLVLDLVALPNNYDTRYPYETTAQ 300
Qy 301 LTRREVYTDPIVFNRETSQGFRCRWLSNLSDISSEVESAVIRSPHLDILSEIEFFYTRAG 360
Db 301 LTRREVYTDPIVFNRETSQGFRCRWLSNLSDISSEVESAVIRSPHLDILSEIEFFYTRAG 360
Qy 361 LPLNTEYLEYVWGHISIKYNT-NASSALERNYGTITSNKIKYDLANKDIFQVRSGLAD 419
Db 361 ITLNNDAYINWNGHTLKYRRRTADSTVTYANYGRITSEK-NSFALEDREDIFEINSTVAN 419
Qy 420 LANYAQAUVGPVYASTFLDKNTGSGVGFYTSKPHTTMOVCTQNTYNTIDIPPENE-P 478
Db 420 LANYAQAUVGPVYASTFLDKNTGSGVGFYTSKPHTTMOVCTQNTYNTIDIPPENE-P 478
Qy 479 LRGYSHRLSHITSYSGFNASSPARYGNLPVFWANTRSDVNTVYSDKITQIPWKAH 538
Db 478 VAESYSHRLSHITSYSGFNASSPARYGNLPVFWANTRSDVNTVYSDKITQIPWKAH 538
Qy 539 TLVSGTIVKPGFTGNTKRTSGPLAYSVKSPLSQRYSRIRYASTTNLRFLVY 598
Db 536 MLVYLGSGVQPGFTGNTKRTSGPLAYSVKSPLSQRYSRIRYASTTNLRFLVY 598
Qy 599 ISGTRIYSINWNTKMGDDLPNTFDLATIGTAFPSNYSLTVGADSPASGGBVYD 658
Db 595 YLGDTTEKRNFKNTMDGASLTYETFKFASFTDFQFRETQDKILLSMGDFSSGQBYD 654
Qy 659 KPELIPVNAFEEEDLDVAKAVNGLFTSKDALQTSVTDVQVQAAANLVECLSDLYP 718
Db 655 RIEFIPVDETEAEQDLAKAVNGLFTSKDALQTSVTDVQVQAAANLVECLSDLYP 714
Qy 719 NEKRLMDVAKKRLVQARNLQDTGPNRINGENGWGTGSGIEVAGDVLFKDRSLRLT 778
Db 715 NEKRLMDVAKKRLVQARNLQDTGPNRINGENGWGTGSGIEVAGDVLFKDRSLRLT 774
Qy 779 SAREIDTETPTLYQOIBESLKPTRYKLGFGIGSSQDLKLRHRANQIVKQVDPN 838
Db 775 GAREIDTETPTLYQOIBESLKPTRYKLGFGIGSSQDLKLRHRANQIVKQVDPN 834
Qy 839 LLPDLVLPVNSCGGIDRCSEQVYDANLANNENGNGWMSDHSFHDITGIDLNENT 898
Db 835 LLPDLVLPVNSCGGIDRCSEQVYDANLANNENGNGWMSDHSFHDITGIDLNENT 888
Qy 899 GIWVVPKIPPTNGYATLGNLEVEEGLSGETLERAQOQEQQWQDQVARKRGASEKAYYA 958
Db 889 GIWVVPKIPPTNGYATLGNLEVEEGLSGETLERAQOQEQQWQDQVARKRGASEKAYYA 948
Qy 959 AKQADRLPADYQDKLNSGVMSDMLAQNLVQSIPTVYNDALPIPGWNTYSPFELTN 1018
Db 949 SKQADRLPADYQDKLNSGVMSDMLAQNLVQSIPTVYNDALPIPGWNTYSPFELTN 1008
Qy 1019 RLQQAANLYDLRANIPNGDFRNLSDWNATSDVNVQOLSPTSVLVLPNNWSQVSOQFTVQ 1078
Db 1009 RLQQAANLYDLRANIPNGDFRNLSDWNATSDVNVQOLSPTSVLVLPNNWSQVSOQFTVQ 1068
Qy 1079 PNRYVYLVRTARKEGVDGXYIIRGDANQOTETLTFNICDDDTGVLSADQTSYITKTVEFT 1138
Db 1069 PNRYVYLVRTARKEGVDGXYIIRGDANQOTETLTFNICDDDTGVLSADQTSYITKTVEFT 1128
Qy 1139 PSTEQWIDMSFEGVNIIESVELVLEE 1167
Db 1129 PYTDQWIBISETEGTFYIESVELVLEE 1157

```

RESULT 5

```

US-07-915-203-2
; Sequence 2, Application US/07915203
; Patent No. 5359048
; GENERAL INFORMATION:
; APPLICANT: Ohba, Michio
; APPLICANT: Iwahana, Hidenori
; APPLICANT: Sato, Reiichi
; APPLICANT: Suzuki, No. 5359048ukazu
; APPLICANT: Ogiwara, Katsutoshi
; APPLICANT: Sakanaka, Kazunobu
; APPLICANT: Hori, Hidetaki
; APPLICANT: Asano, Shouji
; APPLICANT: Kawauchi, Tadaaki
; TITLE OF INVENTION: No. 5359048el Microorganism and Insecticide
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/915,203
; FILING DATE: 19920723
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M/K 301
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1149 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-915-203-2

```

```

Query Match 60.6%; Score 3662.5; DB 1; Length 1149;
Best Local Similarity 63.8%; Pred. No. 4.6e-260;
Matches 752; Conservative 109; Mismatches 257; Indels 61; Gaps 21;

Qy 1 MSPNNQNEYIILDASSTSVSDNSVRYPLANDOTTLLQNMNYKDYLRHSEGENPELFGNP 60
Db 1 MSPNNQNEYIILDASSTSVSDNSVRYPLANDOTTLLQNMNYKDYLRHSEGENPELFGNP 60
Qy 61 ETFISS-STVQIGIGVGOVLGALGVPPAGQIASFVSFVQGLWPSSTSVSNEMIMKQVE 119
Db 61 GTFISAGDVGIGIDIVSTIISGLGIPVLGVFSILGSLGLWPSNNENWQIFNVRVE 120
Qy 120 DLIDQKITDSVRKTALAGLQGLGDLDVYQKSLKNWLENRNDTRARSVVVTVIALELDF 179
Db 121 ELIDQKITDSVRKTALAGLQGLGDLDVYQKSLKNWLENRNDTRARSVVVTVIALELDF 180
Qy 180 VAKIIPSPAISSQEQEPLLSVYAQAANLHLLLRDASIFGAEWGFTPGBEISFYDQRTTAA 239
Db 181 RTNMGFSQNTYETPLLPYTAQAASLHLLVMDVQIYKGEWGPQNDIDL FYKEQVSYTA 240
Qy 240 QYSDYCVKWTNGLDKLKGNTAASLWLYKHQFREMTELLVLDLVALPNNYDTRYPYETTA 299
Db 241 RYSDHCYQWYNAGLNKRGTKAQWVDYNNFRRENNVMDLVALPNNYDTRYPYETTA 300
Qy 300 QLTREVYTDPIVFNRETSQGFRCRWLSN-----SDI-----SFSEVESAVIRSPHLDIL 349
Db 301 ELTREIFTDV-----GSYVTGQSTLISWYDMIPALPSPFTLEN-LLRKPDPFTLL 352

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 10:10:20 ; Search time 30 Seconds
(without alignments)
3741.850 Million cell updates/sec

Title: US-10-089-678-1
Perfect score: 6044
Sequence: 1 MSPNNQNEYIILDASSSTSV.....MSETGCVFNIESVHLVEEE 1167

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3733.5	61.8	1160	2 I40589	parasporal crystal
2	3123	51.7	1157	1 S49247	parasporal crystal
3	2750.5	45.5	1154	2 S39536	parasporal crystal
4	2547.5	42.1	1228	2 S00873	parasporal crystal
5	2458.5	40.7	1138	2 A48944	parasporal crystal
6	2191	36.3	1156	2 S19306	parasporal crystal
7	2154.5	35.6	1174	2 A42459	parasporal crystal
8	2093.5	34.6	1166	2 S32645	parasporal crystal
9	2064.5	34.2	1155	2 A26513	parasporal crystal
10	2061.5	34.1	1155	2 JD0002	parasporal crystal
11	2052.5	34.0	1155	2 S02134	parasporal crystal
12	2048	33.9	1176	2 JT0241	parasporal crystal
13	2045	33.8	1176	2 JC2219	parasporal crystal
14	2044.5	33.8	1181	2 A41052	parasporal crystal
15	2043	33.8	1156	2 A29125	parasporal crystal
16	2038	33.7	1176	2 S02215	parasporal crystal
17	2037	33.7	1176	2 A22617	parasporal crystal
18	2030.5	33.6	1155	2 I39838	parasporal crystal
19	2024.5	33.5	1171	2 I40572	parasporal crystal
20	2010.5	33.3	1171	2 A37829	parasporal crystal
21	1989.5	32.9	1174	2 S32649	parasporal crystal
22	1984	32.8	1189	2 S00944	parasporal crystal
23	1953.5	32.3	1160	2 S32647	parasporal crystal
24	1944	32.2	1165	2 S11446	parasporal crystal
25	1939	32.1	1178	1 USBSXH	parasporal crystal
26	1932	32.0	1156	2 A29838	parasporal crystal
27	1930.5	31.9	1177	2 A49785	parasporal crystal
28	1923.5	31.8	1176	2 A48970	parasporal crystal
29	1917	31.7	1172	2 S32689	parasporal crystal

30 1868.5 30.9 1136 1 USBS81 parasporal crystal
31 1778.5 29.4 1180 2 I39870 parasporal crystal
32 1772.5 29.3 1180 2 A26858 parasporal crystal
33 1734.5 28.7 719 2 I40590 insecticidal prote
34 1722.5 28.5 719 2 I39815 parasporal crystal
35 1713.5 28.4 719 2 S25383 insecticidal prote
36 1707.5 28.3 719 2 I39814 parasporal crystal
37 1683.5 27.9 934 2 A22798 parasporal crystal
38 1413 23.4 823 2 S04181 parasporal crystal
39 1372.5 21.1 652 2 I39811 parasporal crystal
40 1258.5 20.8 659 2 S10228 parasporal crystal
41 1245 20.6 652 2 A27323 parasporal crystal
42 1132.5 18.7 934 2 B29838 parasporal crystal
43 1116.5 18.5 649 1 JH0261 parasporal crystal
44 1069.5 17.7 655 2 JC7140 protoxin - Bacillu
45 1061 17.6 380 2 B42459 hypothetical prote

ALIGNMENTS

RESULT 1
I40589
parasporal crystal protein cry8Cal - Bacillus thuringiensis
N;Alternate names: parasporal crystal protein cry11
C;Species: Bacillus thuringiensis
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 01-Dec-2000
C;Accession: I40589
R;Sato, R.; Takeuchi, K.; Ogiwara, K.; Minami, M.; Kaji, Y.; Suzuki, N.; Hori, H.; Asan
Curr. Microbiol. 28, 15-19, 1994
A;Title: Cloning, heterologous expression, and localization of a novel crystal protein
A;Reference number: I40589; MUID:94100786; PMID:7764305
A;Accession: I40589
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1160 <RES>
A;Cross-references: EMBL:U04366; NID:G532523; PIDN:AAA21119.1; PID:G532524
C;Superfamily: parasporal crystal protein
C;Keywords: delta-endotoxin

Query Match 61.8%; Score 3733.5; DB 2; Length 1160;
Best Local Similarity 64.2%; Pred. No. 1.6e+208;
Matches 767; Conservative 109; Mismatches 257; Indels 61; Gaps 21;

QY 1 MSPNNQNEYIILDASSSTSVSDNSRYPLANDQTTLQNNYKDYLRMSGENDELFCNP 60
DB 1 MSPNNQNEYIILDASSSTSVSDNSRYPLANDQTTLQNNYKDYLRMSGENDELFCNP 60
QY 61 EYFIS- STVQIGIGVQVGLGALGVFPAGQIASPYSPIVQOLWPSSTVSWEIMKQVE 119
DB 61 GTFISAQDAVGTGIDIVSTIISGLGIPVLGEVFSILGSLGLLWPSNNVWQIFMNRVE 120
QY 120 DLIDQKITSVKRTALAGLQGLDGVYQKSLKNWLENNDTRARSVVTVYALDELDF 179
DB 121 ELIDQKILSVRSRATLADANSRIAYEYQVALDWRKNPHSTRSAALVKERFGNAAIL 180
QY 180 VAKIPSAISGQEVPLLSVYAQANLHLRLDASIFGAEWGPTPGHISTFYDQVTRTA 239
DB 181 RTNMGSGSQNTYETPLPTVYAQAASHLLAWRDVQVIGKMGYPONDILDPYKQSVTA 240
QY 240 QVSDYCVKWNVTGLDKLGTNAASWKYQFREMILLVLDVALFNYDTRTPYIETTA 299
DB 241 RYSDHCQVQWYAGLNKRGTKARQWYDNRFRRENNVWLDVALFNYDARIYPLETNA 300
QY 300 QLTREYVYTPVFNRETSGFCRRMSLN-----SDI-----SFSEVESAVIRSPHLPDIL 349
DB 301 ELTRFETDPV-----GSIVTQSGSTLSISWYDMIPALPSPSTLEN-LLRKPDPFTLL 352
QY 350 SEIEFYTRAGLPLNNT-EYLEYVWGHISIKYKNTNASSALERNYGTITSNKIKYDILANK 408
DB 353 QEIRWYTS---FRQNGTIEYVWGGORLTLVIYGSF--NKYSGVLAGAEDIIPVQON 407
QY 409 DIPQVRSGLADLANYYAQVGVYASFTLLDKNTGSGSGVGFYYSKPHHTNQVCTQNT 468

```
Db 408 DIYRV--VMTYIGRYTNSLLGVNPTP--YFSNNTOK-----TYSKP-----KQFAGGIKT 454
Qy 469 ID---EIPENEPISRGYSHRLSHITSYSPKSNASSPARYGNLPVPAWTHRSADVNTVY 525
Db 455 IDSGBELTYEN---YQSISHRVSYITSPEIKSTGTV--LGVPFPGWTHSSASRNFIY 509
Qy 526 SKTIQIPVVKARTLVSGTIVKGRG--FTGQNLKRTSSGPLAY---TSVSVKSPLSORY 581
Db 510 ATKISQIDPINKASRTSGCAVWNPFGYNGGPMKLSGSGSQVIMLRVATDAKG--ASQRY 568
Qy 582 RARTYASTTNLRFLFTTIS-----GTRIYSINVKNKTKNGDDLTFTNFDLATIGTA-F 633
Db 569 RIRIYASDRACKP--TSSRSPENPATYSASIAVTNTMSTNASUTYSTFEAVBSGPINL 626
Qy 634 TFSNYSDSLTVCADSFASGGEVYVDKFEIPVNAATFEABEEDLDVAKKAVNGIFTSKDAL 693
Db 627 GISGSRFPDISITKEAGANLYIDRIEFPVNTLFEABEEDLDVAKKAVNGIFTKEKDAL 686
Qy 694 QTSVTDYQVNOQANLVECLSDBELYPNEKRMMDVAKKAVNGIFTKEKDAL 753
Db 687 QTSVTDYQVNOQANLVECLSDBELYPNEKRMMDVAKKAVNGIFTKEKDAL 746
Qy 754 GWTGSGTGLEVARGDVLFDORSRLRTSAREIDTETPTLYQOI DESLLKPYTRYKLGFI 813
Db 747 GWTGSGTGLEVARGDVLFDORSRLRTSAREIDTETPTLYQOI DESLLKPYTRYKLGFI 806
Qy 814 GSSQDLKLEIKLRHRANQIVKQNPVNDLLPDVLPVNSCGGIDRCSBQQYVDANLALNNGEN 873
Db 807 GSSQDLKLEIKLRHRANQIVKQNPVNDLLPDVLPVNSCGGIDRCSBQQYVDANLALNNGEN 866
Qy 874 GMSSDSHAFSPHIDTGEIDALNENGTVVVFVKIPPTNGVATIGNLELVEBGLSGETLER 933
Db 867 GMSSDSHAFSPHIDTGEIDALNENGTVVVFVKIPPTNGVATIGNLELVEBGLSGETLER 926
Qy 934 ACOQEQOQODKWARKEGSEKAYAAKQAIIDRLADYODOKLNSGVMSDMLAQNLYQS 993
Db 927 ACOQEQOQODKWARKEGSEKAYAAKQAIIDRLADYODOKLNSGVMSDMLAQNLYQS 986
Qy 994 IPYVYNDALPEIPGKNYTSFTBELTNRLQOANLYDLRINAIPNGDFRNLSDNNATSDVNV 1053
Db 987 IPYVYNDALPEIPGKNYTSFTBELTNRLQOANLYDLRINAIPNGDFRNLSDNNATSDVNV 1046
Qy 1054 QOLSSTSVLIPNNNSQVSQSTQPNRYVLRVTRARKEGVDGVVIRIDGANQTETLTP 1113
Db 1047 QOLSSTSVLIPNNNSQVSQSTQPNRYVLRVTRARKEGVDGVVIRIDGANQTETLTP 1106
Qy 1114 NICDDDTGVLSDADOTSYITKTVEFTPSTEQWIDMSETEGVNIESVELVLEE 1167
Db 1107 NICDDDTGVLSDADOTSYITKTVEFTPSTEQWIDMSETEGVNIESVELVLEE 1160

RESULT 2
S49247
parasporal crystal protein cry9Cai [validated] - Bacillus thuringiensis
N:Alternate names: parasporal crystal protein cryIH
C:Species: Bacillus thuringiensis
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 01-Dec-2000
C:Accession: A59350; S49247
R:Lambert, B.; Buysse, L.; Decock, C.; Janssens, S.; Pienas, C.; Saey, B.; Seurinck, J.; V
Appl. Environ. Microbiol. 62, 80-86, 1996
A:Title: A Bacillus thuringiensis insecticidal crystal protein with a high activity agai
A:Reference number: A59350; MUID:96141404; PMID:8572715
A:Accession: A59350
A:Molecule type: DNA
A:Residues: 1-1157 <LAW>
A:Cross-references: EMBL:237527; NID:9547554; PIDN:CAA85764.1; PID:9547556
A:Experimental source: serovar tolworthi
C:Comment: This parasporal crystal protein, active against corn borer and other insects,
C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 51.7%; Score 3123; DB 1; Length 1157;
```

```
Best Local Similarity 52.3%; Pred. No. 4.4e-173;
Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;
Qy 1 MSPNNQNEYEILDSSSTSVSDNSVRYPLANDQTTTLQNMNYKYLRMSEGENPELFGNP 60
Db 1 MNRNNQNEYEIIDA PHCGCPDSDDDVRYPLASDPNAALQNMNYKYLDQMTDEDYDVSINP 60
Qy 61 ETPFSS--STVQIGIVGQVIGALGVPPAGQIATSPISFIVGQVLPSSSTSVVEMIMKQVE 119
Db 61 SLSISGRDAVQTALTVVGRILGALGVPPSGQIVSFYQFLLNTLMPVNDTAIWEAFNRQVE 120
Qy 120 DLIDQKITDSVRKTAAGLQIGLGDGLDVYQKSLKWLLENRNDTRARSVVVTVVIALLELDF 179
Db 121 ELVNQOITTEFARNQALRLQGLGDSFNVTQORSLQNLADRNDTRNLSSVVRAGFIALDLDF 180
Qy 180 VAKIPSAISQGEVPLLSVYAQAANLHLLLDASIFGAEMGTFPGESITFTYDQVTRTA 239
Db 181 VNAIFLFAVNGQVPLLSVYAQAANLHLLLDASIFGEGMGTFQGEISTYTDRLQELTA 240
Qy 240 QYSDYCVKMYNTGLDKLKGNTAASMLKYHOFREMTLLVLDLVALPPNYDTRYPIETTA 299
Db 241 KYTNYCETMYNTGLDRLGNTESWLRYHOFREMTLLVLDVVALPPYDTRYPIETTSNP 300
Qy 300 QLTREVYTDPIVFNRETSGGFCRRKSLNSDISFSVESAVIRSPHLFDILSELFEFTTTRA 359
Db 301 QLTREVYTDPIVFNPPANVGLCRMGNTPYNTFSELENAFIRPPHLPDLRLNSLTISNR-- 359
Qy 360 GLPLANTLEYLVWGHISIKYKNTNASSALERNYGTITSNKIY---YDLANKDIFQVRSL 416
Db 360 -FPV--SSNFMVMSGHTLRRSYLNSAQVEDSYGLITTTTRATINPGVDGNTNR----IEST 413
Qy 417 GADLANYTAQVGVPPYASFTL--LDKNTGSGVGGFTYSKPHYTMQCTQNTYTDIIPP 474
Db 414 AVDFRSALIGIYGVNRASFPVPGCLFNGITTSPPANG-----CRDLYTDNDELPP 461
Qy 475 ENEPISRGVS--HRLSHITSYSPSKN--ASSPARYGNLPVPAWTHRSADVNTVYSDKITOI 532
Db 462 DE---STGSSHLRSHVTFPFSQTNQAGSINAGSVPTVWTRERDVLDLNTIIPNRTQL 518
Qy 533 PVVKAHTLVSGTIVKGRGFTGCGNILKRTSSGPLAYTSVKSPLSQRYARIRYASTTN 592
Db 519 PLVKASAPVSGTIVKGRGFTGCGNILKRTSSGPLAYTSVKSPLSQRYARIRYASTTN 578
Qy 593 LELPVTISGTRIYSINVKNKTKNGDDLTFTNP---DLATIG---TAFTSNYSDSLTGVA 646
Db 579 FSIIRLGRGVSIGDVRILGSTMGRQELTYESFTREFTTGTGPFNPPFTTQAQELITVNA 638
Qy 647 DSFASGGEVYVDKFEIIPVNAATFEABEEDLDVAKKAVNGIFTSKDALQTSVTDYQVNOQA 706
Db 639 EGVSTGGEYIIDRIEIVPNPAREABEEDLEAKKAVASLFTTRDGLQVNVTDYQVDOQA 698
Qy 707 NLVCLSDBELYPNEKRMMDVAKKAVNGIFTSKDALQTSVTDYQVNOQA 764
Db 699 NLVCLSDBELYPNEKRMMDVAKKAVNGIFTSKDALQTSVTDYQVNOQA 758
Qy 765 EGVSTGGEYIIDRIEIVPNPAREABEEDLEAKKAVASLFTTRDGLQVNVTDYQVDOQA 824
Db 759 EGVSTGGEYIIDRIEIVPNPAREABEEDLEAKKAVASLFTTRDGLQVNVTDYQVDOQA 814
Qy 825 RHRANQIVKQNPVNDLLPDVLPVNSCGGIDRCSBQQYVDANLALNNGENGMSSDSHAPS 884
Db 815 HHEKHLVKNVNDLVSDTYSQSCGINRCDEQHQVDMQDLAEHHPMDCEAAQTHBPS 874
Qy 885 PHIDTGEIDALNENGTVVVFVKIPPTNGVATIGNLELVEBGLSGETLERAAQQEQWQDK 944
Db 875 SYINTGDLNASVDQGIWVVKVTRTDGVTATLGNLELVEBGLSGETLERAAQQEQWQDK 934
Qy 945 MARKGASEKAYAAKQAIIDRLADYODOKLNSGVMSDMLAQNLYQVSIPIVYNDALPE 1004
Db 935 LGRKKAIDRVYLAARQAINHLFVDYQDQQLNPEIIGLAINESASNLVBSISGVISDTLLQ 994
Qy 1005 IPGMNVTSTFELTNRLQOANLYDLRINAIPNGDFRNLSDNNATSDVNVQOLSSTSVLVI 1064
```

995 IPGINYBIYETSLDRLOQAASVLYTSRNAVQNGDFNSGLDSNNTTMDASVQODGNMHLVL 1054

1065 PNWNSQVSQOFTVPQPNRYVLRVTARKEGVDGVIIRIDGANGQTETLTFNICDDDTGVLS 1124

1055 SHWDAQVYQQLRVNPNCKYVLRVTARVGGDGVYTRDGAHGHQETLTFNACDYVNGTY 1114

1125 ADQTSYITKTVFTPTSPTEQWIDMSHTEGVNIESVELVLEER 1167

1115 VNDSYITEEVVFYFETKHMVVESEGSFYIDSIEFIETOE 1157

RESULT 3

S39536

parasporal crystal protein cry38a1 - Bacillus thuringiensis

N:Alternate names: delta-endotoxin-related protein; parasporal crystal protein cryX

C:Species: Bacillus thuringiensis

C:Date: 07-Oct-1994 #sequence_revision 01-Dec-1995 #text_change 01-Dec-2000

C:Accession: S39536

C:Author: S. Shevelev, A.B.; Svarinsky, M.A.; Karasin, A.I.; Kogan, Y.N.; Chestukhina, G.G.; Stepan

FEB5 Lett. 336, 79-82, 1993

A:Title: Primary structure of the cryX, the novel delta-endotoxin-related gene from Bacil

A:Reference number: S39536; MUID:94085596; PMID:8262221

A:Accession: S39536

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1154 <SHE>

A:Cross-references: EMBL:X75019

C:Superfamily: parasporal crystal protein

C:Keywords: delta-endotoxin

Query Match 45.5%; Score 2750.5; DB 2; Length 1154;

Best Local Similarity 48.7%; Pred. No. 1.9e-151;

Matches 562; Conservative 197; Mismatches 341; Indels 53; Gaps 18;

Qy 41 NYKDYLPKSEGENPELFCNPTETFISSVTGIGVQVILGALGVPPAGQIASFYISVIG 100

Db 29 SYKDYLPKSEGDYIDSYINPGV--RTGLQRTGIDIVAVVVGALGPPVGGILTGFLSTLFG 86

Qy 101 QLWPSSTVSVVMEMIMKQVEDLIDOKITDSVRKTAGLAGLQGLDGLDVYKSLKNWLENRN 160

Db 87 FLWTSNQAVWEARIEQHEELIEQISDOVVRTALDDLGTQIYNNQYVIALKWEERPN 146

Qy 161 DTRARSVVVQYVIALELDFVAKIISPALSGO-----EVPILLSVYAQAANLHLLLLRDASI 215

Db 147 GVRA-NLVQRPELHALFVSSMPSFG-SGPGSRFOAQLLVVYAQAANLHLLLLRDASI 204

Qy 216 FGAWGFTPGHISFTYDQV--TRTAQVSDYCKVYNTGLDKLKTNAASHLKVHOPREM 274

Db 205 YGARWGLRBSOIGMLYFNEQLQTRTDYTHCNVYNNGLAGLRTGSABSLKYHQPREEA 264

Qy 275 TLLVLDVALPNDYTRTYPLETTAQLTREYVTDPIVFNRETSGP---CRRWLSNSDI 330

Db 265 TLMAMDIALPPTNRYPIVNPQLTREYVTDPIVFPSESSLFPELRLRQETSAM 324

Qy 331 SFSVESAVIRSPHLFDILSEIBFYTTAGLPLNNTLEYLVVGHISI--KYKNTNASSAL 388

Db 325 TFSNLENAIISGPHLFDINNLMYTGVSFVHLN-QLIEGWIGHSVTSLLASGPTTVL 383

Qy 389 ERNYGTITSNKKIYDLANKDI PQVRS---LGADLANYYAOVGVYPVASFLLDKNTGSG 445

Db 384 RNNYGSVTS-IVNYPFSDNRDQVQINTSRSHYGLGFQN--APLFGITTAQF----- 430

Qy 446 SVGGFTYSKPHTYMQVCTQNTYNTIDEIP--PENEPLSGYSHRLSHITSYSP-----SK 497

Db 431 -YPGGTYSVTQNALTCQNTYNSIDELPSLPNEPISRSYSHRLSHITSYLHRVLTIDGI 489

Qy 498 NASSPARYGNLPVPAWTHRSADVNTVYSDKITQIPVYKAHTLVSGTIVIMKPGFTGQNI 557

Db 490 NIYS-----GNLPYTWTHRDVLTNTIYADRIQTQLPVKSPFEPAGITVVRGPGFTGDI 545

Qy 558 LKRTSSGPLAYTSVSKVSPLSQRYRARIYASTTNLRFTVITSGTRIYSINNVNMTMKG 617

Db 546 LRRITGVGFTGIRVETAPLQRIREFASTNLPIGIRVGDQVNVYFPGFTMNRGD 605

Qy	618	DLTENTPDLATIGTATTFNFSYSDSLTVGADSPASGEGVYVDKPELIPVNAATPAREBDLDV	677
Db	606	ELRYESFATREFTTDFNFPQOBSLISVFANAFSAGQEVYFDRLEIIPVNPAREAKEDLEA	665
Qy	678	AKKAVNGLFTSKKDALQTSVTDYOVNQAANLVBCLSDLEYPNPKRMKMDAVKCAKRLVQA	737
Db	666	AKCAVASLFRTRDGLQVNVKDIQVQQAANLVSCLSDQEGYDKMLLEAVRAAKRLSRE	725
Qy	738	RNLLODTGFNRIING--ENGWYTGSTGIEVARGDVLFKDRSLRLTSAREIDTFTPTLYYQQ	795
Db	726	RNLLOQDPDFNTINSTENGWKAANGVTISEGGPFYKGRALQALASAR-----ENPTTYIYQK	781
Qy	796	IDBSLLKPYTRYKLGKPIGSSQDLLEIKLIRHRANQIVKQVDPNLLPDVLVWNSCGGIDRC	855
Db	782	VDASELKPYTRYRSDDGFVKSQDLLEIDLIIHHKHVHLVQKVPNDLVSDTYPDSCSGINRC	841
Qy	856	SEQOYVDANLALFNNGE-KGNMSSDSHAFSFHIDTGTCHIDLNENTGTWVVPKIPPTNGYAT	914
Db	842	QEQQMVAQLQLETHHHHPDMCCEAQAQTHFSSYIDTGDLSNSVDDGQWAIIPKVRTTGGYAT	901
Qy	915	LGNLELVVEEPLSGETLERAQQOEOQWQKARKGRASBKAYVAAKQAIIDRLFADYQDQK	974
Db	902	LGNLELVVEGLPSGESLREQRDNTKWSAELGKRAETORVYQDAKQSIHNLFDVYDQDQ	961
Qy	975	LNSGVEMSDMLAAQNLVQSIPYYVNDALPEIPGKNYTSFTFELNRLQQAANLVDLRNAIP	1034
Db	962	LNPEIGHADIMDAQNLVASISDVYSDAVLQIPGINVEIYITELSNRLQQAASLYTSRNAVQ	1021
Qy	1035	NGDFRNLSDWNTSDVNVQOLSDTSLVLIIPMKNNSQVSQOPTQVQNVRYVLRVTRARKEGV	1094
Db	1022	NGDFNNGLOSDWNTAGASVQDGNTHPLVLSHWDQAQVSQQFRVQPNCKYVLRVTRAEKVG	1081
Qy	1095	GDGVYIIRGANOTELTTFNICDDDTGVLASADQTSYITTKTVFTPTPTEQWIDMSETEGV	1154
Db	1082	GDGVYIIRGAAHTETLTTFNACDYDINGTYVTDNTYLTKEVIFYSHTEHMMVEVNETEGA	1141
Qy	1155	FNIESVELVLEE	1167
Db	1142	FHIDSIEFVETEK	1154

RESULT 4

S00873

parasporal crystal protein cryaB - Bacillus thuringiensis subsp. thuringiensis

N:Alternate names: parasporal crystal protein cryaB

C:Species: Bacillus thuringiensis subsp. thuringiensis

C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 01-Dec-2000

C:Accession: S00873

R:Brizzard, B.L.; Whiteley, H.R.

Nucleic Acids Res. 16, 2723-2724, 1988

A:Title: Nucleotide sequence of an additional crystal protein gene cloned from

A:Reference number: S00873; PMID:88203216; PMID:3362680

A:Accession: S00873

A:Molecule type: DNA

A:Residues: 1-1228 <R1>

A:Cross-references: EMBL:X06711; MID:g40264; PIDN:CAA29898.1; PID:g580949

C:Genetics:

A:Gene: cryaB

A:Start codon: TTC

C:Superfamily: parasporal crystal protein

C:Keywords: delta-endotoxin

```

Query Match      42.1%; Score 2547.5; DB 2; Length 1228;
Best Local Similarity 43.8%; Pred. No. 1.3e-139;
Matches 558; Conservative 183; Mismatches 381; Indels 151; Gaps 21;

Qy      1 MSPNNQNEYRILDASSTVSQNSVRYPLANDQTTTLQNNYKYDYLRMSGEENPELPQNP 60
         | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      1 MTSNRKNENIINAVNSHA-----QNDLLPDARIEDSLCIAEG-----NNI 42

Qy      61 ETPFISSTVOTGIGVQVIGALGVPPAGQIASFYPSFVIGQLWSPSSTVSVMEMTKMOVED 120
         | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

Db 43 DPFTSASTVQGINIAGRIILGVLPFAGQQLASFTSFLUGELNPRGR-DQWEIEFLEHVEQ 101
Qy 121 LIDOKITDSVARTALAGIQLGDLGVYQKSLKNWLENNDTRASVVVVTQYIALELDPV 180
Db 102 LINGQITENARTALARIQLGDSFRAQQSLEDWLENRDDARTSVLYTQYIALELDFTL 161
Qy 181 AKIPSPAIQSQEVPLLSVYQAANLHLLLDASIFGAENGFTGCEISTFYDQVTRTAQ 240
Db 162 NAMPLFAIRNQEVPLLMVYQAANLHLLLDASIFGSEFGLTSQBIQRYYERQVTRD 221
Qy 241 YSDYCVKMYNTGLDKLKGNTAASWLKYHQFREMTELLVLDLVALFPNPTDRTYPIETTAQ 300
Db 222 YSDYCVKMYNTGLSLRGTNAASWVRNQFRDITLGVLDLVALFPNPTDRTYPIETTAQ 281
Qy 301 LTRVYTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSPHLDILSIBIFYYTTRAG 360
Db 282 LTRVYTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSPHLDILSIBIFYYTTRAG 337
Qy 361 LPLNTEYLEVWGHISIKYK-----WTNASSALERNVGTITSNKIKYDILANKDIFQVR 414
Db 338 SRWSENTRHMYRGHTIQSRPIGGLNTSTHGANTNTSINPVTLR-----FASRDVYRTE 391
Qy 415 SLGADL--ANYAAQVYGVYASFTLLDKNTGSGVGGFTYKSPHTTMOVCTQNTYNTIDEI 472
Db 392 SYAGVLLWGLVLEPIHGVTVRFTNPQONISDR-GTANYSQPYESPGLQLKDSFT--EL 448
Qy 473 PPE--NEPLSRGYSHRLSHITSYFSKNASSPARYGNLVPFAWTHRSADVNTVYSDKIT 530
Db 449 PPEPTERRPNYSYSHRLSHIGIILQSR-----VNVVYVSWTHRSADRTNITGPNRIT 500
Qy 531 QIPVVKATLVSGTIVTKPGFTGCTGNILKRTSSGGLATYSVSKSPLSQRYARIRVAST 590
Db 501 QIPVVKASELPQGTIVVRGPGFTGDLRRNTGCGPIRVTVNGPLTQRYIRGFYAST 560
Qy 591 TNLRLFTVISTRIYSINVNKTKMGDDLTNTDPLATIGTAFNYSNLSLTGVGADSPA 650
Db 561 VDPDFVSRGTTVNNFRFLTMSSGBELKYNFVRRAFTPTPTTQIQDIIRTSIOGLS 620
Qy 651 SSGEVYVDKFLIPVNAITFEABEDLVAKKAVNGLFTSKKD-ALQTSVTDYQVNNQANLV 709
Db 621 GNGEYVYDKIILIPVNAITFEABEDLVAKKAVNGLFTSKKD-ALQTSVTDYQVNNQANLV 680
Qy 710 ECLSDLEYPNEKMLDAVBAKLVQANLLODTGNRN----- 750
Db 681 ACUSDPEFLDSEKRELLEKVKYAKLSDBERNLLQDPNFTSINKQDPFITNEQSNFTSIHE 740
Qy 751 -GENGTWGSTGIEVAEGDVLFPKDRSLRLTSAREIDTETPTLYLQQIDESLLKQVTRYKL 809
Db 741 QSEHGHWGSENITIQEGNDVFEKNYVTLPGT---FNECYPTLYLQKIGESSEKAYTRYQL 797
Qy 810 KGPTGSSODLEIKILIRHANOIVRNP--DNLLPDVL--PVNSCGGIDRCSEBQYVDANL 865
Db 798 RGYTEDSQDLEIYLIRYNAKHETLDVPGTESLWPLSVESPIGRGCEPNRCA--PHFEWNP 855
Qy 866 ALENNNGG-NMSSDSHAPSHTDGTEDLNENTGIWVWFKIPTNGVATIGNLELVEEG 924
Db 856 DLDCSDRGKCAHSHHPSLDIDVGCTDLHENIGVWVFKIPTQEGHARIGNLEFIEEK 915
Qy 925 PLSCETLERAOOQOQWDMARKRGASAKAYAAKOIDRLFADYQDQKLNSEVMSDM 984
Db 916 PLLGEALSRVKAEBKWDKREKLOLETKRYVYTEAKAVDALFVDSQYDRLOADNTNIGMI 975
Qy 985 LAAQNLVQSIPIVYNDALPEIPGMNYSFTBLTNRLQOANWLYDLRINAIPNGDFRNLSD 1044
Db 976 HAADKLVRHIREAYLSELFPVPGVNAIBEFEELEGHIIITALSIDARNVWKNKDFNGLTC 1035
Qy 1045 WNATSDVNVQSDTSVLVIPNNSOVSQFTVQPNRYVLRVTRARKEGVDGVYIIRDG 1104
Db 1036 WNVKGHVDVQOQSHRSLSLVIPEWBAEVSQAVRVCPGCCGYILRVYAYKEGVEGCVTIHEI 1095
Qy 1105 ANQETELTF----- 1113
Db 1096 ENWTDLKFKNREESVYPTDTGTCDNTAHQGTAGCADACNSRNAGYEDAYEDVTASV 1155

Qy 1114 -----NICDDDTGVLSAD--QTSYITKTVETPTSTBOVMIDMSETBGV 1154
Db 1156 NYKPTYEBETVDRDNHCEYRGVNVNPPVAGVYTKBLEYFPETDTVMISIGETECK 1215
Qy 1155 FNIESVRLVLEER 1167
Db 1216 FIVDSVELLMEE 1228
RESULT 5
A48944
parasporal crystal protein cry7Aal - Bacillus thuringiensis
N;Alternate names: parasporal crystal protein cryIIIC
C;Species: Bacillus thuringiensis
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 01-Dec-2000
C;Accession: A48944
R;Lambert, B.; Hofte, H.; Anny, K.; Janssens, S.; Soetaert, P.; Peferoen, M.
Appl. Environ. Microbiol. 59, 2536-2542, 1992
A;Title: Novel Bacillus thuringiensis insecticidal crystal protein with a silent activi
A;Reference number: A48944; MUID:92384571; PMID:1514800
A;Contents: BtS137J
A;Accession: A48944
A;Status: preliminary
A;Molecule type: DNA; protein
A;Residues: 1-1138 <LAM>
A;Cross-references: GB:M64478; NID:g142760; PIDN:AAA22351.1; PID:g142761
A;Note: sequence extracted from NCBI backbone (NCBIN:112092, NCBIPI:112093)
C;Superfamily: parasporal crystal protein
C;Keywords: delta-endotoxin

Query Match 40.7%; Score 2458.5; DB 2; Length 1138;
Best Local Similarity 43.5%; Pred. No. 1.6e-134;
Matches 514; Conservative 208; Mismatches 400; Indels 59; Gaps 17;

Qy 1 MSPNQNEYELDASSSTVSVDNSVRYPLANDQTTTLQNMNKKDYLRMSSEGNBELFGNP 60
Db 1 MNLNLDGYE-----DSNRTLNNSLNPTQALSPSKNNMYQDPLSITERQPEALASG 55
Qy 61 EFTFISSTVQTGIGIVQVIGALGVFPAGQIASFYSFIVGQLWFSSTVSVMIMKQVSD 120
Db 56 NT-----AINTVSVTGATLSALCPGASPTNFYKLAGLLWPENG-KIWDEPMTVEA 109
Qy 121 LIDOKITDSVARTALAGIQLGDLGVYQKSLKNWLENNDTRASVVVVTQYIALELDPV 180
Db 110 LIDOKIEYVRNKAELGLGSLDKYQKALADWLKGQDPEALISVATFRIIDSLFE 169
Qy 181 AKIPSPAIQSQEVPLLSVYQAANLHLLLDASIFGAENGFTGCEISTFYDQVTRTAQ 240
Db 170 FMSFVKVTEIPLTVYQAANLHLLLDASITLYGDKGFTQNNIBENTYRQKR15E 229
Qy 241 YSDYCVKMYNTGLDKLKGNTAASWLKYHQFREMTELLVLDLVALFPNPTDRTYPIETTAQ 300
Db 230 YSDHCTKWNSGLSLRGTNAASWVRNQFRDITLGVLDLVALFPNPTDRTYPIETTAQ 289
Qy 301 LTRVYTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSPHLDILSIBIFYYTTR 358
Db 290 LTRVYTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSPHLDILSIBIFYYTTR 338
Qy 359 AGLPLNTE-YLEVWGHISIKYKNTNASSALERNVGTITSNKIK--YYDLANKDIFQVRS 415
Db 339 YKAPSHIQDPLFMSAHKVSFKKSESNLYTTGIYKTSYISGAYSFGNDIYRTLA 398
Qy 416 LGADLANYYAQVGVYASFTLLDKNTGSGVGGFTYKSPHTTMOVCTQNTYNTIDEI 475
Db 399 APSVWVYPTQNYGVEQVEFY-----GVKGHVYRGDNKYDL---YDSDIDQLPDD 446
Qy 476 NEPLSRGYSHRLSHITSYFSKNASSPARYGN--LPVPAWTHRSADVNTVYSDKITIP 533
Db 447 GEPTHEKYTHRLCHATAI-----PKSTPDYDNATIPFSWTHRSAEYIRYPNKITKIP 501
Qy 534 VVKAHTLVSGTIVTKPGFTGCTGNILKRTSSGGLATYSVSKSPLSQRYARIRVASTNL 593
Db 501 VVKAHTLVSGTIVTKPGFTGCTGNILKRTSSGGLATYSVSKSPLSQRYARIRVASTNL 593

Query Match	34.6%	Score 2093.5	DB 2	Length 1166
Best Local Similarity	39.7%	Pred. No. 2.7e-113		
Matches	482	Conservative 172	Mismatches 392	Indels 167
Gaps	24			
QY	59	NPETFISSVTQTGIGVGVLGAL-----GVPFAGQIASPYSFVYLGQMPSSSTVSVM	111	
DB	17	NPSEIENAR-NSMFLGVSVSSGLTFPLEAAVPEAGFALGLFDII-----WGALGVQDM	71	
QY	112	EMIMKQVEDLIDQKITDSVRKATLAGLQGLGDGLDVTYQSKLKNWLENRNDTRASVVVTTQ	171	
DB	72	SLFLRQIEQLIRQIEITELERNRATAILTGLSSSYLYVEALREWENDPNNPASQERVTR	131	
QY	172	YIALELDFVAKIPSFALISGQEVPLLSVYQAANLHLLLLDDASIFGAEMFTPGGEISTFY	231	
DB	132	FRLTDDDAIVTGLPTLAIRNLLEVNLSVYTAANLHLSLLRDADVFGERWGLTQANIEDLY	191	
QY	232	DQWVTRTAQSYDCVKVYNTGLDKLKGTAASWLKYHQFREMTLVLVDLVALPNNVDTR	291	
DB	192	TRLTSTNQIEYSDHCARWYNQGLNEIGGISR----RYLDQFQDLTISVLVDLVALPNNVDIR	247	
QY	292	TYPIETTAQUTREVYTDPIVFNRETSGGFCRRWSLNSDISGFSEVASIVIRSPHLFDILSE	351	
DB	248	TYPIPTQSQLTREIYTSFV-----AGNINFLSIANVRAPHLMDPIDR	292	
QY	352	IEFYTTTACGLPANTTEYLETWVGHISIKYKNTNASSALERN---YGT-----	394	
DB	293	IVIYT-----NSVRSTPYWAGHEVISRKTGGQGGNEIRPPLYGVAANAEPVVTIRPTG	345	
QY	395	ITSNKIKYVLDANKDIEQVRSISGLADLANYYAQQVGVFVPSFTLLDKNTGSGSGGFT-YS	453	
DB	346	FTDEQRQWY-RARSRVVSFRSSGQD-----FSLVD-----AVGFLIIPS	383	
QY	454	KPHITTMQVCTQNYNTIDIEIPENEPRLSGYSHRLSHITSYSPSKNASSP--ARYGNLPVF	511	
DB	384	AVSIYRNGFGFNTDITIDIEPIEGTDPFTGYSHRLCHVGFL-----ASSPPTISQYARAPI	438	
QY	512	ANTHRSADVNTVYSDKIQTPIPVKAHTLVSGCTTVIKGPGTGGNIIKRPSSGGLATSV	571	
DB	439	SWTHRSATLNTIAPDVITQIPLYKAFNLHSGATIVKPGPGTGGDIIILERTNVGSGFGDMRV	498	
QY	572	SVKSPISQRYRIRYASTTNLRLFVITISGTRIYSINVNKTMKNGDDLTFNTPDLATIGT	631	
DB	499	NITAPLSQRTVRIRYASTTDLOPYTWINGTTINIGNPSMTDSGDDLOVGRFVAGFTT	558	
QY	632	AFTPSYSDSLTVGADSFASGEYVYVDKPELIPVNATFEABEDLDVAKCAVNGCLFTSKKD	691	
DB	559	PPTFSDANSTFTIGAPGFSPPNVEYIDRIEFPVPAEVTFEAEYDLEKAQKAVNALFTSSNQ	618	
QY	692	-ALQTSYTDVQVNOANLVECLSDLELYPNEKBLMDAWKAEKRLVQARNLLDGTGFNRIN	750	
DB	619	IGLKTVDYTHIDKVSMLVECLSDFCLEDKRELSEKVKHAKRLSDERNLLQDPNFRGIN	678	
QY	751	--GENGWGTGTGEVASEGDVLFKDRSLRLTSAREIDTETPTYLYQQIDBSLKPYYTRYK	808	
DB	579	RQPDGRWGSTDITIQGGDDVFKENYVTLPGTFD---GCYFTYLYQKIDBSKLVYTRYQ	735	
QY	809	LKGFIGSSQDLIEKILIRHRANQIVKNVPDN-----LLPDVLFPVNSCGGIDRCSEQQYVDAN	864	
DB	736	LRGYIEDSQDLIEYLIRYNARHETVWVPVGTGSLWPLSAQSPIGFKCGEPPNRCA--PHLEWN	793	
QY	865	LALENNGENG-NWSSDSHAFSPHIDTGEIIDLNENTGIWVVFKPIPTTNGYATLGNLELVEE	923	
DB	794	PDLDSCSRNKECAHHSHHFLSDIDVGCTDLNEDLGVWVFIKTKTQGHARLGNLEFLIE	853	
QY	924	GPLSGSETLRAQQOEQQQWQDMARKGASEKAYYAAKQAIDRLPADYQDQKLANSGVMSD	983	
DB	854	KPLLGEALARVKARKKWRDCEKLELETWIVYKEAESVDALFVNQYDQLQADWTIAM	913	
QY	984	MLAAQNVLVQSIPVYVNDALPEIPGNNTYSTFELNRLQQAAWNLVDLNAIINGDFRNLGS	1043	
DB	914	IHAADKEVHGISIREAYLPELSVIPGVNAAIIFEELEGRIFTAFSLYDARNVTKNGDFNNGLS	973	

Qy 1044 DNATSDTAVV-QQLSDTSVLVIWNANVSQSQFTVQVNYVYLRTVTKEGVGDEGYVILR 1104
 Db 974 CMNVKGHVDVEBQNNHRSVLVVPWEAEARVQVCPGRCGYILRTVAYKEGVBEGCVTIH 1033
 Qy 1103 DGANOTETLTNFN-----ICDDDTG-----VLSA 1125
 Db 1034 EIENTYDELKFSNCVEEVYPNNTVTCNDVTYANQEBYKGAYTSNRCGYDEAYGNPSPVA 1093
 Qy 1126 DOT-----SYITKTVPSTPTEQVQWIDMSETEGV 1154
 Db 1094 DYPVVEEKAYTGGRRNPCCSNRGYGDYTPLPAGYVTKLEYFPETDKVWIEIGTEGT 1153
 Qy 1155 PNIESVELVLEE 1167
 Db 1154 FIVESVELLMBE 1166

 RESULT 9
 A26513
 parasporeal crystal protein - Bacillus thuringiensis (strain aizawai)
 C:Species: Bacillus thuringiensis
 C:Date: 11-Mar-1988 #sequence_revision 11-Mar-1988 #text_change 01-Dec-2000
 C:Accession: A26513
 R:Oeda, K.; Ohie, K.; Shimizu, M.; Nakamura, K.; Yamamoto, H.; Nakayama, I.; O
 Gene 53, 113-119, 1987
 A:Title: Nucleotide sequence of the insecticidal protein gene of Bacillus thuri
 A:Reference number: A26513; MUID:87248103; PMID:3297927
 C:Superfamily: parasporeal crystal protein
 C:Keywords: delta-endotoxin

 Query Match 34.2%; Score 2064.5; DB 2; Length 1155;
 Best Local Similarity 39.0%; Pred. No. 1.3e-111;
 Matches 476; Conservative 184; Mismatches 387; Indels 173; Gaps 27;

 Qy 53 NPGL-----FGNPET-FISSSTVQTG---IGI---VGQVLCALGVPPAGQIASPYSP 97
 Db 4 NPINECIPYNCLSNPDEVLGGERIETGYTPDIDLSLTQFLSPFGVPGAGPVLGLVDI 63

 Qy 98 IVGQLWPSSTVSVMEMIKQVEBOLIDQKIDTSVRKLTALAGLQGLGDLYQKSLKNWLE 157
 Db 64 IWGIFGPSQ---WDAPLVQIEQLINQRIEFPARNQAISRLGSLNLYQIYAESSPREWEA 119

 Qy 158 NRNDTRASVVVQYIALELDFVAKIPSPAISQGVPLLSVYAQAANLHLLLRDASIFG 217
 Db 120 DPTNPALREBMRIOFNDMNSALTAIPLEPAVQYQVPLLSVYVQAANLHLSVLDRYDSVFG 179

 Qy 218 AEWGTFPGESTYDQVTRTAQYSDYCVKYNVTGLDKLTGNAASLWKLKVPFRREMTLL 277
 Db 180 QRWGFOAATINSRYNLDITLIGHYTHAVRWYNTGLERVMGPPSDRWIRVQFRRELTLT 239

 Qy 278 VLDELALFPNYDRTYPIETTAQLTREVYTDPIVFNRETSGGFCRRMSLNSDISFSFEVES 337
 Db 240 VLDIVSLFPNYDSRTYPIRTVSQILTRHYNPVL-----ENFDGSRALAAQ 285

 Qy 338 AV---IRSPHLFDILSEIFPTTRAGLPLNNTYLYEYWGHSYK-----YKNTNASSALE 389
 Db 286 GIEGSIKSRPHLMDILNSITVT-----DAHRGEY--YWSGHQIMASPVGFGSPGFPTPL- 337

 Qy 390 RNYGTTSNKIKYDYDLANKDIPQVRSIGLADLANYYAQVY--GVPIASFTLLDKNT---G 443
 Db 338 --YGTWGNAAPOORIVAQLQGQGYRVTLSSTL-----YRRPNIGINNQOLSVLDGTGEYAGT 392

 Qy 444 SGSGVGFTYSKPHYTMQVCTQNTYNTIDEIPPENE--PLSRGYSHRLSHITSY--SFSKNA 499
 Db 393 SSNLPSAVYKSGTV-----DSLDEIPQNNVPPRQGFPHRLSHVSMFSGFSNNS 444

 Qy 500 SSPARYGNLPVFAWTHRSADVTNVTYSDKIQIPIPVVKAHTLVSGTIVYKGPFGTGNIILK 559

Db 791 PI-----GKAAHSHHSLDIDVGTDLNEDLGWVIFKIKTQGHARLGNLE 838
Qy 920 LVERGPISGETLRAQOOQOQDMARKRGASEKAYAAKOAIDRLPADYQDOKLNSGV 979
Db 839 FLEEKPLVGEALARKRAEKWKDKREKLEWETNIYVKEAKESVDALPVNSOYDRLQADT 898
Qy 980 EMSDLAAQNLVQSPYVYNDALPEI PCMYTSTFTELNRLOQAANLYDLBNALPBGDPR 1039
Db 899 NIAMIHAADKRVHSIREAYLPSELSVPGVNAAI FEELEGRIPTAPSLYDARNVKNKGDFN 958
Qy 1040 NGLSDMNATSDVNV-QQLSDTSVLVIPNNWQSOQFTVQPNRYRVLRTVATKKEGVGDY 1098
Db 959 NGLSCWNVKGVHVDVEEQNNHRSVLVPEWBEAEVSEVRCRGYILRTVATKKEGVGEC 1018
Qy 1099 VIIRGANQOTTLTFN-----ICDDTGV----- 1122
Db 1019 VTIIENNTDELKFSNCVEBEVYVNNVTVCNDYTATQEEYEGVTSSNRGYDGAYESNS 1078
Qy 1123 -LSADQTS-----YITKTVETPSTEOVWIDMSE 1150
Db 1079 SVPADYASAYBEKAYTDCRRNPNCSNRGYDGYTLPPLAGYVTKLEYFPETDKVWIBIGE 1138
Qy 1151 TEGVFNIESVELLSE 1167
Db 1139 TEGTIVDSVELLAMEE 1155

RESULT 12
JT0241
paraesporal crystal protein - Bacillus thuringiensis (strain aizawai IPL7)
N:Alternate names: 135K insecticidal protein
C:Species: Bacillus thuringiensis
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 01-Dec-2000
C:Accession: JT0241
R:Shimizu, M.; Oshie, K.; Nakamura, K.; Takada, Y.; Oeda, K.; Ohkawa, H.
A:Title: Cloning and expression in Escherichia coli of the 135-kDa insecticidal protein
A:Reference number: JT0241
A:Accession: JT0241
A:Molecule type: DNA
A:Residues: 1-1176 <SHI>
C:Note: b. thuringiensis aizawai IPL7 produces similar 130K and 135K insecticidal proteins
C:Comment: The 135K protein has insecticidal activity against Plutella xylostella larvae
C:Superfamily: paraesporal crystal protein
C:Keywords: delta-endotoxin

Query Match 33.9%; Score 2048; DB 2; Length 1176;
Best Local Similarity 39.3%; Pred. No. 1.2e-110;
Matches 480; Conservative 174; Mismatches 414; Indels 152; Gaps 26;

Qy 53 NPEL-----FGNPET-FISSSTVQTG---IGI---VGQVLGALGVPPAGQIASFYSP 97
Db 4 NPNINECIPYNCLSNPRVEVLGGERIETGTPIDISLSTPQLLSEFVFGAGFVLGLVDI 63
Qy 98 IVGQLWPSSTVSWEMIKQVEDLIDOKITDSVRKTALAGLOGDGLDVYKSLKNWLE 157
Db 64 IWGIFGPSQ----WDAFLVQIEQLNORIEBEPARNOAISRLGLESLNLYQIYVAESFWEA 119
Qy 158 NRNDTRARSVVVTVIALELDPAKIPGFAISGQEVPLLSVYVAQAANLHLLLRDASIFG 217
Db 120 DPTNPALREEMRIQFNDNNSALTALPLFAVQNYQVPLSVYVQAANLHLSVLRDVSVP 179
Qy 218 AEWGTPGEISTFTDQVTRTAQYSDYCVKMYNTGLDKLKGTAASWLKTHQFREMILL 277
Db 180 ORWGPDAAATINSRYNDLTELIGNYTDYAVRWYNTGLERVMGSDRDWRYNQFRELTLT 239
Qy 278 VLIDLALPNDYTRTYPIETTAOLTRVYVTPDIVENRETS-GGFCRRWSLSDISFSVSE 336
Db 240 VLIDLALPNSDYSRPIRTVSQLRBYITNPVLENFDGSGFMGAQRIEQN----- 290
Qy 337 SAVIRSPHLFDILSIEFT-TRAGLPLNNTLEYLWVCHSFK-----YQNTNASSALER 390

291 ---IRQPHLMDILNSITITVDHKG-----FNWMSGHQITASPVGSGPEPAPPLFG 339
Qy 391 NYGTTITSKIKYDYLANKDIFQVRS-----LGADLANYYAQTGYVPYASFTLLDKRT 442
Db 340 MAGNAAPVLT---VSLTGLGIFERTLSSPLRYRIILSGPNNQELFVLDTGTFEPASLTTWL 397
Qy 443 GSGVGGFTYKPKHTTMQVCTQNTYNTIDBIIPENE--PLSRGYSHRLSHITSYSPSKNAS 500
Db 398 PS-----TIYRQGTV-----DSLVDIIPQDINSVPFRAGFSHRLSHVTVMLSQAGAV 444
Qy 501 SPARYGNLPVPAWTHRSADVTNTVYSDKITQIPVVKAHVTLVSGTTVIKPGPGTGGNILKR 560
Db 445 YTLR---APTFSWQHSRSEAFNNIIPSSQITQIPLTKSTNLGSGTSVWKGPGTGGDILRR 501
Qy 561 TSSGPLATYSVSKPSLSQRYRARIYASTTNLRLPVTISGTRISYINNVKNTMNGDDLT 620
Db 502 TSPQGISLIRVNIITAPLSQRYRARIYASTTNLQFHTSIDGRPINQGNPSATMSGNSLQ 561
Qy 621 FNTPLATIGTAPTFSNYSDSLTVGADSFASGGEVYVDKFLIPVNAIPEAESEEDLVAKK 680
Db 562 SGRFATVGTPTFPNFSNGSSVFTLSAHVFNSEVYIDRIEFPVPAEVTFEASVDLERACK 621
Qy 681 AVNGLFTSKOD-ALQTSVTDYQVNAANLVECLSDDELYPNEKRLMDAVKRAKRLVQARN 739
Db 622 AVNELFTSSNQIGLKTQVTHIDQVSNLVECLSDDELKQELSEKVKHAKRLSDERN 681
Qy 740 LLODTGPNRING--ENGWTSGTGIEVAGDVLFKDRSLRLTSAREIDTETPTLYVQOQID 797
Db 682 LLODPNFRGINQLDRGWRGSTDITIQGGDDVFKENYYTLLGTFFD---BCYPTLYLQKID 738
Qy 798 ESKLKYTRYKLGKPGIGSSQDLLEIKLIRHRANQIVKQVDPN---LLPDLVPLVNSCGGID 853
Db 739 ESKLKAYTRYQLRGVIEDSQDLLEIYLRYNAXHETVNVPGTCSLWPLSAQSPICKGEPN 798
Qy 854 RCSEQQYVDANLALENNGENG-NMSSDSHAFSHIDTGEIDLNTGNTGIWVVKIPPTNGY 912
Db 799 RCA--PHLEWNPDLDCSCRDGKCAHSHHFSLDIDVGTDLNEDLGWVIFKIKTQDGH 856
Qy 913 ATLGNLILVERGPLSGETLRAQOOQOQDMARKRGASEKAYAAKOAIDRLPADYQD 972
Db 857 ARLGNLLEFLKPLVGEALARKRAEKWKDKREKLEWETNIYVKEAKESVDALPVNSQY 916
Qy 973 OKLNGSVEMSDMLAAQNLVQSPYVYNDALPEI PCMYTSTFTELNRLOQAANLYDLRNA 1032
Db 917 DQLOADTIAMIHAADKRVHSIREAYLPSELSVPGVNAAI FEELEGRIPTAPSLYDARNV 976
Qy 1033 IPNGDFRNGLSMNATSDVNV-QQLSDTSVLVIPNNWQSOQFTVQPNRYRVLRTVATK 1091
Db 977 IKNGDFNNGLSLWNVKGVHVDVEEQNNHRSVLVPEWBEAEVSEVRCRGYILRTVATK 1036
Qy 1092 EGVGDGYVILRDGANQOTTLTFN-----ICDD----- 1118
Db 1037 EGYGEGCVTHIENNTDELKFSNCVEBEVYVNNVTVCNDYTATQEEYEGVATSRNRGYN 1096
Qy 1119 DTGVLSDQTS-----YITKTVETPSTEOVWID 1147
Db 1097 EAPSVPADYASVYBEKSYTDGRRNPNCSNRGYDGYTLPPLAGYVTKLEYFPETDKVWIE 1156
Qy 1148 MSETGVFNIESVELLSE 1167
Db 1157 IGTEGTIVDSVELLAMEE 1176

RESULT 13
JC2219
paraesporal crystal protein cryIIa - Bacillus thuringiensis
C:Species: Bacillus thuringiensis
C>Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 01-Dec-2000
C:Accession: JC2219
R:Udayasuriyan, V.; Nakamura, A.; Mori, H.; Masaki, H.; Uozumi, T.
A:Title: Cloning of a new cryII(a) gene from Bacillus thuringiensis strain FU-2-7 and
A:Reference number: JC2219; MUID:94289859; PMID:7764972

A:Accession: JC2219
A:Molecule type: DNA
A:Residues: 1-1176 <UDA>
A:Cross-references: DDBJ:D17518; NID:g506190; PIDN:BA04468.1; PID:g535781
C:Genetics:
A:Gene: cryIA(a)
C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 33.8%; Score 2045; DB 2; Length 1176;
Best Local Similarity 39.3%; Pred. No. 1.9e-110;
Matches 479; Conservative 174; Mismatches 415; Indels 152; Gaps 26;

Qy 53 NPGL-----FGNPET-FISSSTVOTG---IGI---VGOVLGALGVPPAGQIASFYSP 97
Db 4 NPKINECIPYCNLSNPEVVLGGRIETGYTPIDISLSLTQFLISEFVGGAGFVLGLIDL 63
Qy 98 IVGQLWPSSTVSVNEMIMKQVEDLIDQKITDSVRKTALAGLQGLDGLDVYQKSLKNWLE 157
Db 64 IWGIVGFSQ----WDAFLVQIEQLISQRIEFARNQALSRLEGLSNLYQIVAEAFREWEA 119
Qy 158 NRNDTRARSVVVTVQYIALELDFVAKIPSAISGOEVLPSVYQAQANLHLRLDASIFG 217
Db 120 DPTNPALREEMRIQFNDMSALTTAIPLLAVQNVQVPLLSVYQAANLHLSLVDLSVFG 179
Qy 218 AEWGFTGEISTFPYDRQVTRTAQSDYCKVKNVNTGLDKLTNAASWLKYHQFREMILL 277
Db 180 QRWGLDVAATINSRYNDLTRLIGTYTDVAVRYNTGLERWGPDSRWYNGQFRELTLT 239
Qy 278 VLDLVALFPNYDRTYPIETTAQTLREYVTDPIVFNRETSGFCRRWSLNSDISFSEVES 337
Db 240 VLDIVSLFPNYDRTYPIETTSQTLREYVTDPIVFNRETSGFCRRWSLNSDISFSEVES 289
Qy 338 AVTSRPHLFOILSIBFTTATAGLPLNNTEVLEYVWGHSHK-----YKNTNASSALERNY 392
Db 290 S-IRSPHLMDILNSITITTDAGH-----GYTWSGHQINASPVGSGPEFTPEL---Y 338
Qy 393 GTTISNKKYDLANKQIFQVRSILGADLANYYAQ--VVGVPVASFLLDKNT----GSGS 446

Qy 913 ATLGNLELVEEGPLSGETLERAOOQOQWODYKMARKGASEKAYYAAKQADRLPADYOD 972
Db 857 ARLGNLEFLBEKPLVGBALARYVGAEKWDRKREKLEWETNIVYKRAKESVDALPVNSOY 916
Qy 973 QKLSNGVEMSDMLAAQNLVOSIPVYNDALPEIPGNVNTSFTELTNRLQQAWNLVDLNA 1032
Db 917 DQLQADTNIAHIAADKRVHSIREAYLPESLVIFGVNAAIFEELEGRISTAFSLYDAENV 976
Qy 1033 IPNGDFRNGLSDMNATSDVNV-QQLSDTSVLVIPNMSQVSOQFTVQNPVYVLAVTARK 1091
Db 977 IKNGDFRNGLSDMNATSDVNV-QQLSDTSVLVIPNMSQVSOQFTVQNPVYVLAVTARK 1036
Qy 1092 EGVGCDGVVITRDGANQOTELTEN-----ICDO-----1118
Db 1037 EGVGCGCVTHIENNTDELKFSNCVBEIYPNNTVTCNDYTVNQEYGGAYTSRNRGN 1096
Qy 1119 DTGVLSDAOTS-----YITKTEFTPTSEQWID 1147
Db 1097 EAPSPADYASVVEEKSYYTDGRRNCPENRGYRDYTPLPVGYVTKELEYFPTDKWIE 1156
Qy 1148 MSRTGVFNIESVELLEE 1167
Db 1157 IGETEGTFIVDSVELLKEE 1176
RESULT 14
A41052
parasporal crystal protein cryAel - Bacillus thuringiensis (strain alesti)
C:Species: Bacillus thuringiensis
C:Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 01-Dec-2000
C:Accession: A41052
R:Lee, C.S.; Aronson, A.I.
J. Bacteriol. 173, 6635-6638, 1991
A:Title: Cloning and analysis of delta-endotoxin genes from Bacillus thuringiensis subs
A:Reference number: A41052; MUID:92011442; PMID:1655719
A:Accession: A41052
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1181 <LEE>
A:Cross-references: GB:M65252; NID:g142874; PIDN:AAA22410.1; PID:g142875
C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 33.8%; Score 2044.5; DB 2; Length 1181;
Best Local Similarity 39.0%; Pred. No. 1.9e-110;
Matches 477; Conservative 190; Mismatches 404; Indels 151; Gaps 29;

Qy 53 NPGL-----FGNPET-FISSSTVOTG---IGI---VGOVLGALGVPPAGQIASFYSP 97
Db 4 NPKINECIPYCNLSNPEVVLGGRIETGYTPIDISLSLTQFLISEFVGGAGFVLGLIDL 63
Qy 98 IVGQLWPSSTVSVNEMIMKQVEDLIDQKITDSVRKTALAGLQGLDGLDVYQKSLKNWLE 157
Db 64 IWGIVGFSQ----WDAFLVQIEQLISQRIEFARNQALSRLEGLSNLYQIVAEAFREWEA 119
Qy 158 NRNDTRARSVVVTVQYIALELDFVAKIPSAISGOEVLPSVYQAQANLHLRLDASIFG 217
Db 120 DPTNPALREEMRIQFNDMSALTTAIPLLAVQNVQVPLLSVYQAANLHLSLVDLSVFG 179
Qy 218 AEWGFTGEISTFPYDRQVTRTAQSDYCKVKNVNTGLDKLTNAASWLKYHQFREMILL 277
Db 180 QRWGLDVAATINSRYNDLTRLIGTYTDVAVRYNTGLERWGPDSRWYNGQFRELTLT 239
Qy 278 VLDLVALFPNYDRTYPIETTAQTLREYVTDPIVFNRETSGFCRRWSLNSDISFSEVES 337
Db 240 VLDIVSLFPNYDRTYPIETTSQTLREYVTDPIVFNRETSGFCRRWSLNSDISFSEVES 289
Qy 338 AVTSRPHLFOILSIBFTTATAGLPLNNTEVLEYVWGHSHK-----YKNTNASSALERNY 392
Db 290 S-IRSPHLMDILNSITITTDAGH-----GYTWSGHQINASPVGSGPEFTPEL---Y 338
Qy 393 GTTISNKKYDLANKQIFQVRSILGADLANYYAQ--VVGVPVASFLLDKNT----GSGS 446

Db 339 GTMGNAAPQORIVAOAGQGVKTLTSL---STFRNPFIIGINNQRLSVLDGTEFAVGSSN 395
QY 447 VGGFTYSKPHITMQVCTQNTYNTIDIPPE--NEPLSRGYSHRLSHITSY--SFSKNASSP 502
Db 396 LPSAYLRSGTV-----DSLDIIPQDNNVPPRQGFHLSHVSMFRSGFSNYSVI 447
QY 503 ARYGNLPVPAWTHRADVTNTVYSDKITQIPIVVKARTLVSGTIVTKGPGFTGGNLLKRTS 562
Db 448 IR--APWFSMIHRSABFNIIIPSSQITQIPLTKSTNLGSGTSVVGKPGFTGGDILRRTS 504
QY 563 SGPLATSVSVKSPSQRVARIYASTNLRLFTVTSIGTRYISVNVKTNKGGDILTFN 622
Db 505 FQGISLTAVNTAPLSQRVIRVASTNLTQFHTSIDGRPNQNFSAITWSSGGNLSQG 564
QY 623 TFDLATIGTAFTHSNGSSVFTLSAHVFNNGEVYIDRIEFVPAEVTFAEYDLERAQAV 682
Db 565 SPRTVGFTTTPNFNSGSSVFTLSAHVFNNGEVYIDRIEFVPAEVTFAEYDLERAQAV 624
QY 683 NGLFTSKKD-ALQTSVTDYQVNAANLVECLSDLEYLPNEKMLMDAVKAGRLVQARNLL 741
Db 625 NALFTSSNQIGLKTVDYHIDQVSNLVECLSDLEYLPNEKMLMDAVKAGRLVQARNLL 684
QY 742 QDTGNRIN--GENGTGSGTIEVAEGDVLKORSRLTSAREIDTETPTLYQQIDBS 799
Db 685 QDPNFRGINRQDRGWRGSDTITIQGGDDVFKENYVTLPGTDP---ECYPTLYQKIDES 741
QY 800 LKPYTRYKLGKFGICSSODLEIKLRHRANQIVKQNP--DNLLPDLV--PVNSCGGIDRC 855
Db 742 KUKATRYQLRGYIIDSQLEIYLIRYNAKHETVNVPGTSLWPLSFSSICKCGEPNRC 801
QY 856 SBOQYVDANLALENNGENG--NMSSDSHAFSHIDTGEIDLNTGIVVVFKIPTTNGVAT 914
Db 802 A--PHLEWNPDLDCSRDGEKCAHSHHPSLDIDVGDIDLNBGLVGVVIFKIKTQDGHAR 859
QY 915 LGNLELVREGPLSGETLRAQQOQOQWQMARKKRGASEKAYYAKQALDRLFADYQDOK 974
Db 860 LGNLELVREGPLSGETLRAQQOQOQWQMARKKRGASEKAYYAKQALDRLFADYQDOK 919
QY 975 LNSGVMSDMLAAQMLVQSIPIVYNDALPEIQMNVYTSFELTNLQQAANLYDLRNAP 1034
Db 920 LQADTNIAHIHTADKRVHRIQAYLPESLVIQVNAIGIFEELEGRIFTAYSILYDARNV 979
QY 1035 NGDFRNLGSDNATSDVNV--QQLSDTSVLVIENNSQVSSQFTVQPNRYVLRVATARKG 1093
Db 980 NGDFRNLGSDNATSDVNV--QQLSDTSVLVIENNSQVSSQFTVQPNRYVLRVATARKG 1039
QY 1094 VGGYVILRDGANQTELTFTNIC-----DDD 1119
Db 1040 YGEGCVTHIEINIDELKFSNCVBEVYPNNVTVCNEYTANQEBVGGATSCNRYDET 1099
QY 1120 TG---VLSADQTS-----YITKTVFTPTTEQVW 1145
Db 1100 YGSNYSVPADYASVVEEKAYTDGRRENPCSNRGYDTPPLPAGVYTKQLEYFPETDKW 1159
QY 1146 IDMSFEGVNTIESVELVLEE 1167
Db 1160 IBIGETEGTFIVDSVELFLME 1181

RESULT 15
A29125
parasporal crystal protein Bc2 - Bacillus thuringiensis subsp. kurstaki (strain HD-1)
C:Species: Bacillus thuringiensis subsp. kurstaki
C:Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 01-Dec-2000
C:Accession: A29125
R:Fischhoff, D.A.; Bowdish, K.S.; Perlak, F.J.; Marrone, P.G.; McCormick, S.M.; Niedermeyer, J.
A:Title: Insect tolerant transgenic tomato plants.
A:Reference number: A29125
A:Accession: A29125
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-1156 <FIS>

C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 33.8%; Score 2043; DB 2; Length 1156;
Best Local Similarity 39.1%; Pred. No. 2,2e-110;
Matches 476; Conservative 185; Mismatches 389; Indels 168; Gaps 30;

QY 53 NPEL-----FGNPEP-TFSSSTVQTG---IGI-----VGQVLGALGVPPAQOIAFYSP 97
Db 4 NPINIECIPYCNLSNPREVVLGGERIETGYTPIIDISLSTQFLLSFVAGFVLGLVDI 63
QY 98 IVGQLMPSSTVSVWEMIKQVEDLLDQKTDSTVRKLTALAGLQGLDGLDVOYKSLKNWLE 157
Db 64 IWGIFGPGQ---WDAPLVQIEQLNQRIEERPAQOIALRLEGLNLYQIYAESFWEA 119
QY 158 NRNDTRARSVVVTVYIALELDFVAKIPSPAISQGVPLLSVYAQAANLHLLLRDASIFG 217
Db 120 DPTPALREEMRIQPNMNSALTALPLFAVQNYQVPLLSVYQAANLHLSVLRDVSFVG 179
QY 218 ABWGTTPGEISTFYDQRTVTRTAQYSDYCVKYNWNTGLDKLGTNAASLWKVQFRREMTLL 277
Db 180 QRMGFDAAATINSRYNDLTRELIGNYTHAVRWNTGLERVMGPDSDRDWIRYQFRRLTTL 239
QY 278 VLDLVALPNTYDTRTPPIETTAQLTREVVTDPIVFNRETSGGFCRRWSLNSDISPSEVES 337
Db 240 VLDIVSLFPNYSRTPITVTSQLTREIYTNFLENFD--GSF--RGSAGQ-----IEG 289
QY 338 AVIRSPHPLFDILSEIEFTYTRAGLPLNTEYLEYVWGHSHK-----YKVNASSALLERNY 392
Db 290 S-IRSPHMLDILNSITIT---DAHREY--YWSGHQIMASPVGSPGEPFTPL---Y 338
QY 393 GTITSNKIKYDVLANKDIPQVRSGLADLANYYAQVY--GVVYASFTLLDKNT-----GSGS 446
Db 339 GTMGNAAPQORIVAOAGQGVKTLTSLTSL---YRPFNIGINNQQLSVLSDGTEFAYGTSN 395
QY 447 VGGFTYSKPHITMQVCTQNTYNTIDIPPE--NEPLSRGYSHRLSHITSY--SFSKNASSP 502
Db 396 LPSAYLRSGTV-----DSLDIIPQDNNVPPRQGFHLSHVSMFRSGFSNYSVI 447
QY 503 ARYGNLPVPAWTHRADVTNTVYSDKITQIPIVVKARTLVSGTIVTKGPGFTGGNLLKRTS 562
Db 448 IR--APWFSMIHRSABFNIIIPSSQITQIPLTKSTNLGSGTSVVGKPGFTGGDILRRTS 504
QY 563 SGPLATSVSVKSPSQRVARIYASTNLRLFTVTSIGTRYISVNVKTNKGGDILTFN 622
Db 505 FQGISLTAVNTAPLSQRVIRVASTNLTQFHTSIDGRPNQNFSAITWSSGGNLSQG 564
QY 623 TFDLATIGTAFTHSNGSSVFTLSAHVFNNGEVYIDRIEFVPAEVTFAEYDLERAQAV 682
Db 565 SPRTVGFTTTPNFNSGSSVFTLSAHVFNNGEVYIDRIEFVPAEVTFAEYDLERAQAV 624
QY 683 NGLFTSKKD-ALQTSVTDYQVNAANLVECLSDLEYLPNEKMLMDAVKAGRLVQARNLL 741
Db 625 NALFTSSNQIGLKTVDYHIDQVSNLVECLSDLEYLPNEKMLMDAVKAGRLVQARNLL 684
QY 742 QDTGNRIN--GENGTGSGTIEVAEGDVLKORSRLTSAREIDTETPTLYQQIDBS 799
Db 685 QDPNFRGINRQDRGWRGSDTITIQGGDDVFKENYVTLPGTDP---ECYPTLYQKIDES 741
QY 800 LKPYTRYKLGKFGICSSODLEIKLRHRANQIVKQNP--DNLLPDLV--PVNSCGGIDRC 855
Db 742 KUKATRYQLRGYIIDSQLEIYLIRYNAKHETVNVPGTSLWPLSFSSICKCGEPNRC 801
QY 856 SBOQYVDANLALENNGENG--NMSSDSHAFSHIDTGEIDLNTGIVVVFKIPTTNGVAT 914
Db 802 A--PHLEWNPDLDCSRDGEKCAHSHHPSLDIDVGDIDLNBGLVGVVIFKIKTQDGHAR 859
QY 915 LGNLELVREGPLSGETLRAQQOQOQWQMARKKRGASEKAYYAKQALDRLFADYQDOK 974
Db 860 LGNLELVREGPLSGETLRAQQOQOQWQMARKKRGASEKAYYAKQALDRLFADYQDOK 919
QY 975 LNSGVMSDMLAAQMLVQSIPIVYNDALPEIQMNVYTSFELTNLQQAANLYDLRNAP 1034
Db 920 LQADTNIAHIHTADKRVHRIQAYLPESLVIQVNAIGIFEELEGRIFTAYSILYDARNV 979
QY 1035 NGDFRNLGSDNATSDVNV--QQLSDTSVLVIENNSQVSSQFTVQPNRYVLRVATARKG 1093
Db 980 NGDFRNLGSDNATSDVNV--QQLSDTSVLVIENNSQVSSQFTVQPNRYVLRVATARKG 1039
QY 1094 VGGYVILRDGANQTELTFTNIC-----DDD 1119
Db 1040 YGEGCVTHIEINIDELKFSNCVBEVYPNNVTVCNEYTANQEBVGGATSCNRYDET 1099
QY 1120 TG---VLSADQTS-----YITKTVFTPTTEQVW 1145
Db 1100 YGSNYSVPADYASVVEEKAYTDGRRENPCSNRGYDTPPLPAGVYTKQLEYFPETDKW 1159
QY 1146 IDMSFEGVNTIESVELVLEE 1167
Db 1160 IBIGETEGTFIVDSVELFLME 1181


```
Db      899  TNIAIHADKRVHSIREAYLPESLVIPGVNAIPEELBGRIFTAPSLYDARNVIKNGDP 958
Qy      1039  RNLSDNWNATSDNVV-QOLSDTSVLVTPNWNNSQVSQQFTVOPNTRYVLRVTARKEGVGDG 1097
Db      959  NGLSCNWNKGVHDVVEBQNNHRSVLAVPWEAEVSEQVRVCPGRGYILRVTYAYKEGYGEG 1018
Qy      1098  YVIRDGANCTSTLTEN-----ICDDDTGV----- 1122
Db      1019  CVTIHEIENNTDELKFSNCVBEVYPNNVTTCNDYTATQEEYEGYTTSRNRYDGAYESN 1078
Qy      1123  --LSADQTS-----YITKTVBFTPSTEQVWIDMS 1149
Db      1079  SSVPADYASAEKAYTDGRDNPCESNRGYGYTLPAGVYTKLEVPETDKVWIEIG 1138
Qy      1150  ETEGVFNIESVELVLEE 1167
Db      1139  ETEGTFIVDSVELLMEE 1156
```

Search completed: June 21, 2004, 10:13:10
Job time : 37 secs

Db 1 MSPNNQNEYIIDATPSTSVSSDSNRYPFAFNEPTDALQNMNYKDYLRKSGGENSELPFGNP 60
Qy 61 ETPFSSSTVGTGIGVGNLALGVPAGQIASFVPIVGLWPSSTVSVMIMKQVBD 120
Db 61 ETPFSSSTVGTGIGVGNLALGVPAGQIASFVPIVGLWPSSTVSVMIMKQVBD 120
Qy 121 LIDQKITDSVRKTLAGLQGLDGLVYQVSKWLNENRNDTRARSVVVQYIALDLDFV 180
Db 121 LVDQKIEKYVKDALKALGKGNALDYYQVSLDLENRNDTRARSVVVQYIALDLDFV 180
Qy 181 AKTSPFALSGOEVPLLSVYAAANLHLLLDASIPGAEWGFTEGISTFYDROVYTAQ 240
Db 181 SSIPFVAVSGHEVLLVYAAANLHLLLDASIPGAEWGFTEGISTFYDROVYTAQ 240
Qy 241 YSDYCVKMYNGLDKLKTNAASLKYHQFREMTHLVLDVALFPNYDRTYPIETTAQ 300
Db 241 YSDYCVKMYNGLDKLKTNAASLKYHQFREMTHLVLDVALFPNYDRTYPIETTAQ 300
Qy 301 LTREYVTDPIVFNRETSGFCRRWSLNSDISPSVESAVIRSPHLPDILSIEPIYTRAG 360
Db 301 LTRDVTDPDIAFNIVTSGFCRRWSLNSDISPSVESAVIRSPHLPDILSIEPIYTRAG 360
Qy 361 LPLNTEVLEYVWCHSLKYKNT-NASSALEBNYCTITSNKLKYVDLANKOIFQVRSIGAD 419
Db 361 ITLNDAYINYSNGHTLKYRRTADSTVYTYNGRITSEK-NSPALSDRDLFEINSTVAN 419
Qy 420 LANYAQQVYGVYASFTLLDKNTGSGVGGFTYKPHPTMQVCTQNTYNTIDEIPNE-P 478
Db 420 LANYAQQVYGVYASFTLLDKNTGSGVGGFTYKPHPTMQVCTQNTYNTIDEIPNE-P 478
Qy 478 LSRGYSHRLSHITSYSPSKNASSPARYGNLVPFAWTHRSADVTNVTYSDKITQIPVKAH 538
Db 478 VAESYSHRLSHITSYSPSKNASSPARYGNLVPFAWTHRSADVTNVTYSDKITQIPVKAH 538
Qy 539 TLVSGTIVKPGPTGGNKLKRTSGPLAYTSVSKPLSRVRYRARIYASTTNLRFPVT 598
Db 539 TLVSGTIVKPGPTGGNKLKRTSGPLAYTSVSKPLSRVRYRARIYASTTNLRFPVT 598
Qy 598 MLYGGSVVQVGGFTGGDLKRTSPSLGTFVAVTNGSLRVRIRIYASTTDFE-FTL 594
Db 598 MLYGGSVVQVGGFTGGDLKRTSPSLGTFVAVTNGSLRVRIRIYASTTDFE-FTL 594
Qy 599 ISGTPIYSINVKTMKGGDDLTFTNFDLATIGTAFPTSNYSDSLTVGADSPASGEVYVD 658
Db 599 ISGTPIYSINVKTMKGGDDLTFTNFDLATIGTAFPTSNYSDSLTVGADSPASGEVYVD 658
Qy 654 YLGTIEKVRPNKTMNGASLTYETFKFASPIITDQPRETQDKILLSNGDFSSQEVYID 654
Db 654 YLGTIEKVRPNKTMNGASLTYETFKFASPIITDQPRETQDKILLSNGDFSSQEVYID 654
Qy 718 KPELIPVNAATEBEDLDVAKAVNGLFTSKKALQTSVTDYQVQNAANLVECLSDILYP 718
Db 718 KPELIPVNAATEBEDLDVAKAVNGLFTSKKALQTSVTDYQVQNAANLVECLSDILYP 718
Qy 714 RIEPIPVDETYEABQDLEAAKAVNALFTNTKQGLRPGVTDYEVQNAANLVECLSDILYP 714
Db 714 RIEPIPVDETYEABQDLEAAKAVNALFTNTKQGLRPGVTDYEVQNAANLVECLSDILYP 714
Qy 778 NEKGLMDAVKAEKLVQARNLQDTGFRNGENGTGSGIYVAGDVLFPKDRSLRT 778
Db 778 NEKGLMDAVKAEKLVQARNLQDTGFRNGENGTGSGIYVAGDVLFPKDRSLRT 778
Qy 774 NEKGLMDAVKAEKLVQARNLQDTGFRNGENGTGSGIYVAGDVLFPKDRSLRT 774
Db 774 NEKGLMDAVKAEKLVQARNLQDTGFRNGENGTGSGIYVAGDVLFPKDRSLRT 774
Qy 838 SAREIDTETPTLYLQIDSLKPYTRYKLGPGSSQDLLEIKLIRHRANQIVKRVNPDN 838
Db 838 SAREIDTETPTLYLQIDSLKPYTRYKLGPGSSQDLLEIKLIRHRANQIVKRVNPDN 838
Qy 834 GARIDTETPTLYLQIDSLKPYTRYKLGPGSSQDLLEIKLIRHRANQIVKRVNPDN 834
Db 834 GARIDTETPTLYLQIDSLKPYTRYKLGPGSSQDLLEIKLIRHRANQIVKRVNPDN 834
Qy 898 LLPDLVFNKCGGIDRCSBOQYVDANLALNNGENGNMSSSHAPSFHIDTGEIDLMENT 898
Db 898 LLPDLVFNKCGGIDRCSBOQYVDANLALNNGENGNMSSSHAPSFHIDTGEIDLMENT 898
Qy 888 LLPDLVFNKCGGIDRCSBOQYVDANLALNNGENGNMSSSHAPSFHIDTGEIDLMENT 888
Db 888 LLPDLVFNKCGGIDRCSBOQYVDANLALNNGENGNMSSSHAPSFHIDTGEIDLMENT 888
Qy 958 GIWVVPKIPPTNGYATLNLVREGLPSGFTLERAQOEOOQWDMKARKGASEKAYYA 958
Db 958 GIWVVPKIPPTNGYATLNLVREGLPSGFTLERAQOEOOQWDMKARKGASEKAYYA 958
Qy 948 GIWVVPKIPPTNGYATLNLVREGLPSGFTLERAQOEOOQWDMKARKGASEKAYYA 948
Db 948 GIWVVPKIPPTNGYATLNLVREGLPSGFTLERAQOEOOQWDMKARKGASEKAYYA 948
Qy 1018 AKQADRLFADYQDQKLSGVEMSDMLAAQNLVQSIPIVYNDALPEIPGMNYTSFTELTN 1018
Db 1018 AKQADRLFADYQDQKLSGVEMSDMLAAQNLVQSIPIVYNDALPEIPGMNYTSFTELTN 1018
Qy 949 SKQAVDLRYADYQDQKLSGVEMSDMLAAQNLVQSIPIVYNDALPEIPGMNYTSFTELTN 1008
Db 949 SKQAVDLRYADYQDQKLSGVEMSDMLAAQNLVQSIPIVYNDALPEIPGMNYTSFTELTN 1008
Qy 1019 RLQQAANLYDLRNAIPNGDFRNLGSLDMNATSDVNVQQLSDTSVLVIPNNQSVSQPFTVQ 1078
Db 1019 RLQQAANLYDLRNAIPNGDFRNLGSLDMNATSDVNVQQLSDTSVLVIPNNQSVSQPFTVQ 1078
Qy 1009 RLQQAANLYDLRNAIPNGDFRNLGSLDMNATSDVNVQQLSDTSVLVIPNNQSVSQPFTVQ 1068
Db 1009 RLQQAANLYDLRNAIPNGDFRNLGSLDMNATSDVNVQQLSDTSVLVIPNNQSVSQPFTVQ 1068
Qy 1079 PNYTVLVYATKRGVGGYVYIDGACNOTETLAFNICDDTGVLSADQTSYITKTVFT 1138
Db 1079 PNYTVLVYATKRGVGGYVYIDGACNOTETLAFNICDDTGVLSADQTSYITKTVFT 1138
Qy 1069 PNQRYVLVYATKRGVGGYVYIDGACNOTETLAFNICDDTGVLSADQTSYITKTVFT 1128
Db 1069 PNQRYVLVYATKRGVGGYVYIDGACNOTETLAFNICDDTGVLSADQTSYITKTVFT 1128

Qy 1139 PSTQVWIDMSBTRGVFNISVELVLBBE 1167
Db 1129 PYTDQWIEISBTRGVFNISVELVLBBE 1157

RESULT 2

C8CA BACTP
ID C8CA BACTP STANDARD; PRT; 1160 AA.
AC Q45706;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Pesticidal crystal cryc8a (insecticidal delta-endotoxin
DE CrvIIIc(a)) (Crystalline entomocidal protoxin) (130 kDa crystal
protein).
GN CRY8CA OR CRVIIIc(A) OR CRYIII.
OS Bacillus thuringiensis (subsp. japonensis).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=128936;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Bubui.
RX MEDLINE=94100786; PubMed=7764305;
RA Sato R., Takeuchi K., Ogiwara K., Minami M., Kaji Y., Suzuki N.,
Hori H., Asano S., Ohba M., Iwahana H.;
RT "Cloning, heterologous expression, and localization of a novel crystal
protein gene from Bacillus thuringiensis serovar japonensis strain
bubui toxic to scarabaeid insects.",
RL Curr. Microbiol. 28:15-19(1994).
RN (2)
RP SEQUENCE OF 1-14 AND 56-64, AND CHARACTERIZATION.
RX MEDLINE=94259659; PubMed=8200856;
RA Hori H., Suzuki N., Ogiwara K., Himejima M., Indrasith L.S.,
Minami M., Asano S., Sato R., Ohba M., Iwahana H.;
RT "Characterization of larvicidal toxin protein from Bacillus
thuringiensis serovar japonensis strain Bubui specific for
scarabaeid beetles.",
RL J. Appl. Bacteriol. 76:307-313(1994).
CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
EPITHELIAL CELLS OF INSECTS. ACTIVE ON VARIOUS SCARABAEID BEETLES
SUCH AS ANOMALA CUPREA, A. RUPOCUPREA AND POPILLIA JAPONICA.
CC -1- DEVELOPMENTAL STAGE: The crystal protein is produced during
sporulation and is accumulated both as an inclusion and as part of
the spore coat.
CC -1- MISCELLANEOUS: Toxic segment of the protein is located in the N-
terminus.
CC -1- SIMILARITY: Belongs to the delta endotoxin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC EMBL; U04366; AAA21119.1; -
DR PIR; I40589; I40589.
DR HSP; P07130; 1DLG.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_N.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal_bind_like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin; 1.
DR Pfam; PF03945; endotoxin_N; 1.
DR Toxin; Sporulation.
SQ SEQUENCE 1160 AA; 130425 MW; C16C3D912BEB8751 CRC64;

Query Match 61.8%; Score 3733.5; DB 1; Length 1160;
Best Local Similarity 64.2%; Pred. No. 4.4e-205;
Matches 767; Conservative 109; Mismatches 257; Indels 61; Gaps 21;

Qy	1	MSPNQNEVEIL	DA	SS	TS	VS	DN	SR	YPL	AND	QT	TT	LQ	NN	YK	DL	RM	SE	GN	PE	L	GN	P	60	
Db	1	MSPNQNEVEI	IDL	SP	TS	VS	DN	SR	YPL	AND	QT	NT	LQ	NN	YK	DL	KW	TE	SN	AE	L	SR	P	60	
Qy	61	ETPIS-	ST	QT	GI	GV	QV	LG	AL	GP	FA	QI	AS	IF	SF	IV	QV	LP	SP	TS	VS	WEM	K	119	
Db	61	GTPISQA	DV	GT	GI	DV	ST	II	SG	IL	GV	FS	IL	GS	IL	GL	LP	SN	EN	W	QI	F	N	120	
Qy	120	DLIDOKIT	DS	VR	K	T	AL	AG	LG	D	GL	DV	YQ	KS	L	K	W	LE	N	R	N	D	T	179	
Db	121	ELIDOKIL	DS	VR	GA	I	DL	NS	IA	VE	YQ	NA	LE	D	W	R	K	N	P	H	S	T	R	180	
Qy	180	VAKIP	SP	AI	SQ	EV	PL	SV	Y	AQ	AN	L	H	L	L	R	D	AS	I	F	GA	E	W	G	239
Db	181	RTNMG	S	F	SQ	T	N	E	T	P	L	P	T	Y	AQ	AS	L	H	L	V	M	R	DV	Q	240
Qy	240	OYSDY	C	V	K	T	N	T	GL	D	K	L	G	T	NA	S	W	L	K	Y	H	O	P	R	299
Db	241	RYSDHC	Q	W	M	T	N	AG	L	K	R	G	T	GA	K	Q	W	D	Y	N	R	E	M	N	300
Qy	300	QLTRE	V	T	D	P	I	V	N	R	E	T	S	G	F	C	R	W	S	L	N	---	S	D	349
Db	301	ELTRE	I	P	T	D	V	-----	G	S	V	T	C	S	T	L	S	W	D	M	I	P	A	A	352
Qy	350	SEIEF	Y	T	R	A	G	L	P	L	N	T	-	E	Y	L	E	Y	W	G	H	S	I	K	408
Db	353	QEIR	M	W	T	S	---	F	R	O	M	T	I	E	Y	N	Y	W	G	Q	R	L	T	S	407
Qy	409	DIQV	R	S	L	G	A	D	L	A	N	Y	AQ	V	Y	G	P	V	A	S	P	T	L	D	468
Db	408	DIYR	V	-	V	W	T	I	G	R	Y	N	S	L	G	V	N	P	V	T	-	Y	F	S	454
Qy	469	ID---	E	I	P	P	E	N	E	P	L	S	R	G	Y	S	H	L	S	I	T	S	P	S	525
Db	455	IDSGE	L	T	Y	E	N	---	Y	Q	S	Y	S	H	R	V	S	I	T	S	F	E	I	K	509
Qy	526	SDKIT	P	I	P	V	K	A	N	T	L	V	S	G	T	T	V	I	K	G	P	-	F	T	581
Db	510	ATKISQ	I	P	I	N	K	A	S	T	S	G	A	V	N	P	O	E	G	L	Y	G	G	P	568
Qy	582	RARIR	A	S	T	T	N	L	R	L	F	V	T	S	---	G	T	R	I	V	S	I	N	N	633
Db	569	RIRIR	A	S	D	R	A	G	K	P	-	T	I	S	R	S	P	E	N	P	A	T	Y	S	626
Qy	634	TFSN	S	D	L	T	V	G	A	D	S	P	A	S	G	E	V	Y	D	K	F	E	L	I	693
Db	627	GISGS	R	T	F	D	I	S	I	T	K	B	A	G	A	N	L	Y	I	D	R	E	I	P	686
Qy	694	QTSV	T	D	Y	Q	V	N	Q	A	N	L	V	E	C	L	S	D	E	L	P	N	E	K	753
Db	687	QTSV	T	D	Y	Q	V	N	Q	A	N	L	I	E	C	L	S	D	E	L	P	N	E	K	746
Qy	754	GWTG	S	T	G	I	E	V	A	E	G	D	V	L	F	K	D	S	R	L	T	S	A	R	813
Db	747	GWTG	S	T	G	I	E	V	E	G	D</														

```

Qy 1054 QQLSDTSVLVLPWNNSQVSOQFTQPNRYRVLVLTARKEGVGGYVIIRGANGQTETLTP 1111
Db 1047 QQLSDTSVLVLPWNNSQVSOQFTQPNRYRVLVLTARKEGVGGYVIIRGANGQTETLTP 1106

Qy 1114 NICDDDTGVLSDAQTSYITKTVETFTSTEQVWIDMSSETGVFNIESVELVLEE 1167
Db 1107 NICDDDTGVLSDAQTSYITKTVETFTSTEQVWIDMSSETGVFNIESVELVLEE 1160

RESULT 3
C8BA_BACUK STANDARD; PRT; 1169 AA.
AC Q45705;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pesticidal crystal protein cry8Aa (Insecticidal delta-endotoxin
DE CryvIIIa) (Crycaline entomocidal protoxin) (134 kDa crystal
DE protein).
GN CRY8BA OR CRYVIIIB(A) OR 50C(B).
OS Bacillus thuringiensis (subsp. kumamotoensis).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=132267;
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=NRRL B-18746 / PS50C;
RA Michaela T.E., Poncetrada L., Narva K.R.;
RT "Process for controlling scarab pests with Bacillus thuringiensis
isolates.";
RL Patent number W09315206, 05-AUG-1993.
CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
EPITHELIAL CELLS OF INSECTS. ACTIVE ON VARIOUS SCARABID BEETLES.
CC -1- DEVELOPMENTAL STAGE: The crystal protein is produced during
sporulation and is accumulated both as an inclusion and as part of
the spore coat.
CC -1- MISCELLANEOUS: Toxic segment of the protein is located in the N-
terminus.
CC -1- SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/).
CC or send an email to license@isb-sib.ch.
CC
CC EMBL; U04365; AAA21118.1; -.
CC HSPP; P07130; 1DLC.
CC InterPro; IPR001178; Endotoxin.
CC InterPro; IPR005638; endotoxin_C.
CC InterPro; IPR005639; endotoxin_N.
CC InterPro; IPR008979; Gal_bind_like.
CC Pfam; PF00555; endotoxin_1.
CC Pfam; PF03944; endotoxin_C; 1.
CC Pfam; PF03945; endotoxin_N; 1.
CC Toxin; Sporulation.
KW SEQUENCE 1169 AA; 133543 MW; 22EPCF5BD699909 CRC64;

Query Match 52.2%; Score 3152; DB 1; Length 1169;
Best Local Similarity 54.0%; Pred. No. 6.4e-172;
Matches 653; Conservative 179; Mismatches 294; Indels 84; Gaps 25;

Qy 1 MSPNNQNEYILDASSSTVSVDNSVRYPLANDQTTTLONNYYKDYLPWSEGENPELPGNP 60
Db 1 MSPNNQNEYIIDAFTSTVSNDNSRYPFANBPTNALQNDYDKYLKMSAGNVSEYPCGP 60

Qy 61 ETFTLS-SSTVQTGTGIGVQVILGALGVFPAGQIAFSFYFVQLWPSSTVSVMWIMKQVE 119
Db 61 EVFLSEQDAVKAAADIIVKGLTTLGVGFVGVISLYTLQILDLWPSKQKQWELFMEQVE 120

Qy 120 DLIDQKITDSVRKTLAQLQGLDGLVDVYQKSLKNWLENRNDTRARSVVVQYIALELDF 179

```

Db 121 ELINQKTAAYARNKALSELGIGNNYQLYLTALBEWKENPNSRALRDVNRFLDLSLP 180
Qy 180 VAKIPSPAIQOEVPFLSVYAQAANLHLLLRDASIFGAERGFTPEISTFYDQVTRTA 239
Db 181 TOYMSPFVTFVFPFLVYVYMAANLHLLLRDASIFGEEGLSTSTINNYNRQMKLTA 240
Qy 240 QYSDYCVKWTGDLKLGKTAASWLKTHQFRRENTLVLVLVALPFDYDTRTPPIETTA 299
Db 241 EYSDHCWKVYETGLAKLGSSAKOWIDYNQFRRENTLVLVDVALPSNYDTRTPPLATTA 300
Qy 300 QLTREVTYDPIVENRETSGGFCRRWSLNSDLSFSEVRSVAIVRSPLDILSIBETPTTRA 359
Db 301 QLTREVTYDPIGADVDPNIG---SW-YDKAFSFEIEKAIRPPVFIYGLTYTKR 356
Qy 360 GLPLNTEYLSYVWGHISYKYNKTNASSALERNYGT-----ITSNKIKYDILANKDIPQVR 414
Db 357 S--PTSDDRYMYWAGHQISYKHGTSSTFTQWYGTNQLQSTN----PDPNTVDYKTL 410
Qy 415 SLGADLAN-----YAAQVYGVPIYASFTLLDKNTGSGSGVGGFTYKXPHMTMQVCTQNYTI 469
Db 411 SNGAVLLDIVPGYTYTTFGMPETEFPMVNLNTRKT--LTY-KP-ASKOIIDRTRDSE 466
Qy 470 DEIPPE--NEPLSRGYSRLSHITSYFSKNASSPARVGNLPPFAWTHRSADVNTVYSD 527
Db 467 LELPETSQPNYESYSHRLGHIT-FYSSSTST-----YVPVFSWTHRSADLTNTVKS 520
Qy 528 KITQIPVVKARTLVSGTTVIKPGPTGGNLIKRTSSGPLAYTSVSKSPLSQRYPARIRY 587
Db 521 BITQIPGKSSITIGRTVIKRGYTGCDLVALTR--IGSCFQPMIPESQFRIRIRY 578
Qy 588 AS--TTNLRLP-VTISGTRIVSINVNT-MNKGD-DLFTNTPDLATIGTAPTSNYSDSL 642
Db 579 ASNETSYSLYGLNOSGT-----LKNQTYSNKXENDLTND-----PKYISYPRVI 625
Qy 643 TVGADS-----PASGEVVV-DKPELIPVNAITFEAEEDLDVAKAVNGLFTSKDOL 693
Db 626 SVNASSNTQRLSIGITQTNLFIIDRIEIPVDIYEATDLEAAKAVNALFTYKGL 685
Qy 694 QTSVDYQVNOANLVECLSDLEYPNEKRMMLADVAKRLVQARNLLQDTGCFNRINGEN 753
Db 696 QGVTDYEVNOANLVECLSDLEYPNEKRLLDVAREAKRLSEARNLLQDPDFQBSINGEN 745
Qy 754 QHTSGTGEVARGDVLKDRSRLSAREIDTETPTLYLQOIDEKSLKPYRYKLKGPFI 813
Db 746 GWTASTGTEVIRGDAVFGKRYLLEPCAREIDTETPTLYLQKVEGVLKPYRYKLGPV 805
Qy 814 GSSQDLKILRRANOIVKNPVDMILLPDVLPVNSCGGIDRCSEQQYVDANLALNNGEN 873
Db 806 GSSQGLEIYTIHQTNRIKVNVPDILLPDVPPVNDGRINRSEQKYNSRLEVENR--- 862
Qy 874 GNMSSDSHAFSFDITGEIDLNENTGIWVVPKIPPTNGYATIGNLLEVEGPLSGETLER 933
Db 863 ---SCEAHEFSIPIDTGEIDYENAGIWWGFKITDPGEYATIGNLLEVEGPLSGDALER 919
Qy 934 AQOQEQOQWAKRKGSEKAYAAKQAIIDLPAQYQOKLNSGVENSMDLAQNLVQS 993
Db 920 LQKESQWQKIQMTRRREBETDRYMAKQAVDLADYADYQQLNPVNEITDLTAAQDLIQS 979
Qy 994 IPVYNDALPEIPGNNYTSFTLNRLOQAMNLYDLRNALPNGDFRANGLSDNNATSDVNV 1053
Db 980 IPVYNEPPEIPGNNYTFTELTLRLOQANGLYDORNAIPNGDYRNLSNNTSYGVNV 1039
Qy 1054 QQLSDPSTVLVI PNWNSQVSOQFTVPQNYRYLVLRVTRARKEGVGGVYVIRDGANQETLTF 1113
Db 1040 QQINHTSVLVI PNWNEQVSOKEFTVPQNYRYLVLRVTRARKEGVGGVYVIRDGANQETLTF 1099
Qy 1114 NTCDDDTGVLSDQTS-----YIT-----KTUETPSTEQVWIDMSSETGVNI 1157
Db 1100 SASDYDTNGHYDQASNTNGYNTNSVYMIKPAIRSKTVDISSYTNQMWIBISETGTFYI 1159
Qy 1158 ESVELVLEE 1167
Db 1160 ESVELVLDVE 1169

RESULT 4

C9CA_BACTO
ID_C9CA_BACTO STANDARD; PRT; 1157 AA.
AC Q45733;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pesticidal crystal protein cry9Ca (insecticidal delta-endotoxin
DE CryIXC(a)) (Crystalline entomocidal protoxin). (130 kDa crystal
DE protein).
GN CRY9CA OR CRYIXC(A).
OS Bacillus thuringiensis (subsp. tolworthi).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI TaxID=1442;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=BT02618A;
RX MEDLINE=96141404; PubMed=8572715;
RA Lambert B., Buysse L., Decock C., Jansens S., Piens C., Saey B.,
RT Seurinck J., Van Audenhove K., Van Rie J., Van Vliet A., Peferoen M.;
RT "A Bacillus thuringiensis insecticidal crystal protein with a high
RL activity against members of the family Noctuidae.";
CC Appl. Environ. Microbiol. 62:80-86(1996).
CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF LEPIDOPTERA LARVAE. HAS A FAIRLY BROAD
CC SPECTRUM OF ACTIVITY AGAINST MEMBERS OF THE FAMILIES PYRALIDAE,
CC PLUTELLIDAE, SPRINGIDAE, AND NOCTUIDAE. IT IS THE FIRST
CC INSECTICIDAL CRYSTAL PROTEIN WITH ACTIVITY AGAINST CUTWORMS. NO
CC ACTIVITY IS OBSERVED AGAINST SOME BEETLES, SUCH AS COLORADO POTATO
CC BEETLE.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, Z37527; CAA85764.1; --
CC PIR, A59350; S49247.
CC HSP, P07130; IDIC.
CC InterPro; IPR001178; Endotoxin.
CC InterPro; IPR005638; endotoxin_C.
CC InterPro; IPR005639; endotoxin_N.
CC InterPro; IPR008979; Gal_bind_Like.
CC Pfam; PF00555; endotoxin_1.
CC Pfam; PF03944; endotoxin_C; 1.
CC Pfam; PF03945; endotoxin_N; 1.
CC Toxin; Sporulation.
CC KW
SQ SEQUENCE 1157 AA; 129775 MW; C364391BF7PDB8A CRC64;

Query Match 51.7%; Score 3123; DB 1; Length 1157;
Best Local Similarity 52.3%; Pred. No. 2.8e-170;
Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;
Qy 1 MSPNNQNEYIILDASSSTSVSDNSVRYPLANDQTTTLNNNYKDYLRMSGEPNPLFGNP 60
Db 1 MNRNNQNEYIILDAPHCPCSDDDVRYPLASDPNALQNNYKDYLTQMTDEDYDTSYINP 60
Qy 61 ETPFIS-STVQIGIGVQVLCALGVPPAGQIASFYSPVIGQLWPSSTVSWEMIMKQVE 119
Db 61 SLSISGRDAVQALTAVGRILGALGVPPSQIVSPFYQFLTLNLPVNDTAIMEAFMRQVE 120
Qy 120 DLDQKIDTSVRKTKALAGLQGLDGLDVYQKSLKNLENRNDTRARSVVVTVYIALELDF 179

```
Db 121 ELVNOQITLPEFARNOALRLQGLGDSFNVYORSQNLMLADNRDNRNLGVSVPRAQFIALDLDF 180
Qy 180 VAKIPSPAIQSQEVPLISVYAQAANLHLLLRDASIFGAEMGFTPGBISTIFYRQVTRTA 239
Db 181 VNAIPFAVNGQVPLISVYAQAANLHLLLRDASIFGAEMGFTPGBISTIFYRQVTRTA 240
Qy 240 QYSDYCVKWNVTGLDKLKGNAASWLKXHOFRREMTLLVLDVALFPNYTRTYPIETTA 299
Db 241 KYNICYETWNTGLDRLGRTNTESWLXHOFRREMTLLVLDVALFPNYTRTYPIETTSNP 300
Qy 300 QLTREYVTDIVNRETFSGFCRWSLNSDISSEVESAVIRSPHLPDILSEIEFYTTTA 359
Db 301 QLTREYVTDIVNPPANVGLCRWGTNPNTSELENAPIRPHLPDLRLSLTISNR- 359
Qy 360 GLPLNLEYLYVGHISIKYKNTNASSALERNYGTITSNKIKY---YDLANKDIPQVRSI 416
Db 360 -FPV-SSNFMDYNSGHTLRSYLNDSAVQSDSGLIITTTATINPGVDGTR-----IEST 413
Qy 417 GADLANIYAQVGYVPYASPTL--LDKNTGSGVGGFTYSKPHHTMQVCTQNYNTIDRPP 474
Db 414 AVDFRSALIGIYGNRASFPVGGLPNGTTPSPANGG-----CRDLYDTNDELPL 461
Qy 475 ENPFLSGYS-HRLSHITSYSPKX-ASSPARYGNLPVPAWTHRSADVTYVYSDKITOI 532
Db 462 DE---STGSSTHRLSHVTFPSQTNQAGSIANAGSVPTTYVTRDVLNNTITPNRITOL 518
Qy 533 PVVKAHTLVSGTIVIKGPGFTGNNILKRTSGGLAYTSVSKPSLQRYARIRYASTTN 592
Db 519 PLVNASPVSGTIVIKGPGFTGGILRTTNGFTGLRVTVNSPLTQOYLRLRVFASTGN 578
Qy 593 LRLPVTISGTRIYSINVKTMKRGDDLTFTNP---DLATIG---TAPTFSNYSDSLTVGA 646
Db 579 PSIRVLGGVSGIDVRLGTMNRQELTYESFFTRPTTTPGPNPPTFTQAQELTVNA 638
Qy 647 DSFASGSEVYDFKELLIPVNAITEAEEDLDVAKAVNGLPTSKDALQTSVTDYVNOQA 706
Db 639 EGVSTGSEYVIDRIEIVPNVPAEAEEDLEAAKAVASLPTTRTRDGLQVNVTDYQVDOA 698
Qy 707 NLVECLSDRLYPNEKRLMDAVKAKRLVOARNLLQDTGNRRING--ENGWGTSGTGEVA 764
Db 699 NLVSLCLSDQEGHDKMLRAVRAKRLSRERNLLQDPDFNTINSTENGKASNGVTIS 758
Qy 765 EGDVLPDRSLRLTSAREIDTETPTLYLQOIDBSLKLKPYRYLKGFIGSSQDLKLI 824
Db 759 EGGPFKGRALQLASAR---ENYPTIYQKVDASVLKPYRYRLDGFVKSSQDLKLI 814
Qy 825 RHRANQIVKXVPDNLDPVLPVNSCGGIDRCSEQOYDAMLALNNGENGNWSSDSHAFS 884
Db 815 HHKVLVLRVXVPDNLVSDTYSKSCGGINRCDEQHQVDMQLDAEHHPMDCCAAQTHEFS 874
Qy 885 FHIDTGEIDLNENTGIWVWFKIPPTNGVATLGNLELVEBGLSGTETLERAQOQOQWQDK 944
Db 875 SYINTGDNASVDGCIWVLKVRTTIDGATLGNLELVEVGLPSGESLEREQDNAKWNAE 934
Qy 945 MARKGASAKYAAKQAIIDRLPADYQDOKLNSGVEMSDMLAAQNLVQSIPIVYNDALPE 1004
Db 935 LGRKRAEIDRVYLAQAQAINHLFVDYQDQQLNBPGLABINEASNLVSIISGVYSDTLQ 994
Qy 1005 IPGNYTSTFELTNRLQANLILYDLRNAIPNGDFRNLGSDWNATSDVNVQQLSDTSVLVI 1064
Db 995 IPGNYEITYELSDRLQAQASYLYTSRAVQNGDFNSGLSDSNNTMTDASVQDQGNHFLVL 1054
Qy 1065 PNWNSQVSOQPTVQPNRYLVTRARKEGVDGVYIIRDGANOTETLTFNICDDDTGVLIS 1124
Db 1055 SHWDAQSOQLRVNPNCKYLVTRARKVGGDGVYITRDGAHGETLTFNACDIDVNGTY 1114
Qy 1125 ADQTSYITKTVEPTSTPEQWIDKSETGVFNIESVELVLEE 1167
Db 1115 VNDNSYITEEVVFPETKMAVEVSESGSFYIDSIEPIETOE 1157
```

RESULT 5

```
C9EA_BACTA
ID C9EA_BACTA STANDARD; PRT; 1150 AA.
AC Q9ZNL9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pesticidal crystal protein cry9EA (Insecticidal delta-endotoxin
DB CRYIXE(a) (Crystalline entomocidal protoxin) (130 kDa crystal
DB protein).
GN CRY9EA OR CRYIXE(A).
OS Bacillus thuringiensis (subsp. aizawai).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1433;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SSK-10;
RA Midoh N., Oyama K.;
RT "Bacillus thuringiensis cry gene for insecticidal crystal protein.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases
CC -/- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF INSECTS.
CC -/- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -/- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -/- SIMILARITY: Belongs to the delta endotoxin family.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.eib.ch/announce/
CC or send an email to license@sib.ch).
CC -----
DR EMBL; AB011496; BAA34308.1; -
DR HSSP; P07130; IDLC.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005636; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal_bind_Like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
DR Toxin; Sporulation.
KW Toxin; Sporulation.
SQ SEQUENCE 1150 AA; 129895 NW; 7D6AB93D6EDC97EB CRC64;

Query Match 46.7%; Score 2825.5; DB 1; Length 1150;
Best local similarity 48.6%; Pred. No. 2.6e-153;
Matches 574; Conservative 197; Mismatches 364; Indels 47; Gaps 17;

Qy 1 MSPNNONEYTELDASSSTSVSDNSVRYPLANDQTTTLQNNMYKDYLRMSEGE-----NP 54
Db 1 MNRNPNNEYELIDAPYCGCPSSDDVRPPLASDPNNAAPQNNYKEYLYQYDGYTGSLLNP 60
Qy 55 ELFGNPTFTSSSTVQTGIGVQVGLGALGVPPAGQIASFYSPVIGQWLPSTSVSWEMI 114
Db 61 NLSINPR-----DVLQTGINIVGRILGFLGVPFAGQLVTFYTFLLNQLWPLNDNAWBAF 115
Qy 115 MKQVEDLIDQKITDSVRKTAGLQGLDGLDVQKSLKWLKNLENDTRASSVVTOTYA 174
Db 116 MAQTEELIDQKISAQVVRNALDDLTGLHDHYEYEAALAEWLERNPANGARA-MLVTRFEN 174
Qy 175 LELOFVAKIPSPFATS-QGE---VPLLSVYAQAANLHLLLRDASIFGAEMGFTPGBISTIF 230
Db 175 LATAFVTRMPSFGTGPSSQSDAVALLTYAQAANLHLLLRDASIFGAEMGFTPGBISTIF 234
Qy 231 YDRQVTRTAQYSDYCVKWNVTGLDKLKGNAASWLKXHOFRREMTLLVLDVALFPNYDT 290
Db 235 PNAQOQERITRYTNHCVEYTRNGLEDVRGNTESMLNTHFRFRREMTLMAMDVLVALFPYVNV 294
Qy 291 RTYPIETTAQLTRVYTDPIVFNRETSQGFRCRWSLNSDISSEVESAVIRSPHLPDILS 350
```

```
Db 295 ROYNGANPQLTREIYTDPIVYVPPANQICRRWGNPNYVTFSELENAPFIPPHLPERLN 354
Qy 351 EIEFYTTAGIPLANTTEVLEWVGHISIKYKNTNASSALERNYGTITSNKIKYDLANKDI 410
Db 355 RLTIISRNRYTPTTNS-FLDWSGHITLQSHANNPTTYSYGTQTSN-TRUFNTTN-GA 411
Qy 411 FOVRSLGADLANYYAQVYGVYASFLLDKNTGSGVGGFTYSKPHPTMQVCTQNYMID 470
Db 412 RAIDSRAENFGNLYANLGV--SSLNIPP--TG-----VMSEITNAANTCRQDLTTE 460
Qy 471 HIPPENBELSGYSHRLSHITSYSFSKNASSP-ARYGNLPVPFWATHRSADVNTVYSKI 529
Db 461 ELPLENNP-----NLLSHVTPLRENTTQGGPLATLGFVPTVMTREVDFTNTTADRI 515
Qy 530 TOIPVVKATLVSGTIVTKGPGTCGNILKRTSSGPLAYTSVSKPSQVRARIRVAS 589
Db 516 TOLPFWKASEIGGTVVKGPGTGDILRRDGGAVGTIRANVWAPLQQRIRLRVAS 575
Qy 590 TTN--LRLFVTISGTRIYSINVNKTMRKGGDITFNTFPOLATIGTAFTEFSNDSLTVGAD 647
Db 576 TTSFVNLFPVNSAA--GFTLPSTMAQNSLTYESFNTLEVTHTIRFSQSDTLRLNIP 632
Qy 648 SPASGCVYVOKFELIPVNAFPEAREEDLVAKAVNGLFTSKDALQTSVTDYQVNOAN 707
Db 633 PSISQGVVVDKLEIVPINTREAREEDLEDAKAVASLFTTRDGLQVNVTDYQVDOAN 692
Qy 708 LVECLSDRLYNEKEMLDANKAKRLVQARNLQDTCFNIRNG--ENGWGTSGTGEVAE 765
Db 693 LVCLSDSOYGHDKOMLEAVRAAKRLGRNLLQDPDPNEINSTEENGWAKASNGVTISE 752
Qy 766 GDVLFKRLSLRTSAREIDTETPTLYLQQIDESLLKPYTKLKGFGTSSQDLKILIR 825
Db 753 GGPFFKGRALQASAR----ENYPTIYQKVDASTLKPVTYKLDGPFVQSSQDLIDLII 808
Qy 826 HRANOIVKVPDNLPLDVLVNSCGGIDRCSEQQVVDANLALENGNGENGNSSDSHAFSP 885
Db 809 HKHVLVKNVPDNLVSDTYSDCSGCINRCBHQHVQVQDABDHPKDCCEAAQTHFESS 868
Qy 886 HDTGEIDLNENTGIVWVFKIPTTNGYATLGNLELVESGSLGETLRAQQOQOQWQDKM 945
Db 859 YHIGDLNASVQGLVWVQLVTRTDGYATLGNLELVESGSLGETLRAQQOQOQWQDKM 928
Qy 946 ARKGASBKAYAAKQALDRLPADYQDQKLSGVMSDLAAQNLVQSPVYVNDALPEI 1005
Db 929 GKRAETDRIYQDAQAINHLFVDYQDQQLSPEVGMADIIDAQNLIASISDVYSDAVLQI 988
Qy 1006 PGQNTSTFELTNRLQOAWNLVDLNAIPNGDFRGLSDWATSVDVYVQQLSDTSVYVIP 1065
Db 989 PGINTEMYTELNLRLQOASLYTSRNVVQNGDFNSGLDSWNTTDTAVVQDGNMHLVLS 1048
Qy 1066 NWSVSOQFTVQPNRYVRLVYRTAKBVGDCGVYIIRDCANOTETLTENICDDDTGVLSA 1125
Db 1049 HMDAQVSOQFVQPNCKYVRLVYRTAKVNGDGVYIYQDGNHRELTITFNACDYVNGTHV 1108
Qy 1126 DQTSYITTKTEVTPSTEQWIDMSGTGVPNIESVELVEE 1167
Db 1109 NDNSYITKELVFPYKTEHMMWVEVSEGTGYFIDSTIEFTQOE 1150

RESULT 6
C9DA_BACTP
ID_C9DA_BACTP STANDARD; PRT; 1169 AA.
AC O06014:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Pesticidal crystal protein cry9Da (Insecticidal delta-endotoxin
DE CryIXD(a)) (Crystalline entomocidal protoxin) (132 kDa crystal
DE protein).
GN CRY9DA OR CRYIXD(A).
OS Bacillus thuringiensis (subsp. japonensis).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
```

```
OX NCBI_TaxID=128936;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N141;
RA Asano S.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF INSECTS.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; D85560; BAAL1948.1; -.
CC HSSP; P07130; IDUC.
CC InterPro; IPR001178; Endotoxin.
CC InterPro; IPR005638; endotoxin.
CC InterPro; IPR005639; endotoxin N.
CC InterPro; IPR008979; Gal bind_like.
CC Pfam; PF00555; endotoxin; 1.
CC Pfam; PF03944; endotoxin_C; 1.
CC Pfam; PF03945; endotoxin_N; 1.
CC Toxin; Sporulation.
CC KW Toxin; Sporulation.
CC SQ SEQUENCE 1169 AA; 132228 MW; 659AB25729DSE9 CRC64;

Query Match 44.9%; Score 2715; DB 1; Length 1169;
Best Local Similarity 46.9%; Pred. No. 5.3e-147; Indels 46; Gaps 17;
Matches 558; Conservative 204; Mismatches 383;

Qy 1 MSPNQNEYEILDASSSTSVSDNSVRYPLANDQTTTLQNMNYKDYLRMSEGENPELFGNP 60
Db 1 MNRNQNQEYEVIDA PHCCPADDDVVKYPLTDDPNAGLQNMNYKEYLQTYGDDYDPLNP 60
Qy 61 EPTIS-STVTQIGIVQVIGALGVPPAGQIASFYFVQVQLWSPSTSVVEMINKQVE 119
Db 61 NLSVSGKDVIVQVINGIVGLLSFPFGFPSSQWVTYTYLLNSLWPDENSVDAMFERVE 120
Qy 120 DLIDOKITDSVRKTAGLAGLQIGDGLDVYKSLKWLNRNDTRARSVVVTVQIALELDF 179
Db 121 ELIDQKISRAVGRALDDLTLGLQYNYNLVFEALDEWLNRPGARA-SLVSRQFNILDSLF 179
Qy 180 VAKIPSPAISG-----QEVPLLSVYAQAANLHLLLRDASIFGABWGFTPGISTPYDRQ 234
Db 180 TOPMPSFG-SGPGSQNYATILLPVYAQAANLHLLLRDADIVGARWGLNQTQIDQPHSRQ 238
Qy 235 VTRTAQYSDYCVKNTNTGLDKLGNAAASLWKHQFRRMTLLLDLVALFPNYDTRYP 294
Db 239 QSLTQTYTHCVTAYNDCGLAELRGTTAESWPKYQYRREMTLTAMDVALFPYNYLRQYP 298
Qy 295 IETTAQLTRVYTDPIVN--RETSGFCRRWSLN-----SDISFSEVSASVRSRPHLFDI 348
Db 299 DGTNPQLTRVYTDPIAFDPLEQPTQLCRSWINPAPFNHNLNFSVLNSLRPPLPER 358
Qy 349 LSEIEF---YTTAGLPLNNTLEYLVWVGHISIKYKNTNASSALERNYGTITSNKIKYDYL 405
Db 359 LSNLQLVNYQTNGSA-----WRGSRVRYVHLHSSIIQKSYGLSLDSPVGANINV 408
Qy 406 ANKIDIFQVRSAGADLANYYAQVYGVYASFLLDKNTGSGVGGFTYSKPHPTMQVCTQN 465
Db 409 QNNDIYQIISQVSNPASPVGSSYSVMDTNFYL-----SSQVSGISGYTQQGIAPVCLQQ 463
Qy 466 YNTIDEIP---PENEPLSRGYSHRLSHITSYSFSKNAS-SPARY-CNLVPFWATHRSADV 520
```


Db 464 RNSTDDELPSLNPEGD-IIRNYSHRLSHITQYRFOATOSGSPSTVSANLPTCVWTHRDVQL 522
Qy 521 TNTVYSDKTIQIPVWKAHTLVSGTIVKGPFGTCGNILKRTSSGPLAYTQSVSVKSPLSQR 580
Db 523 DNTITANQITQLPLVRAVELSSGATVVGKPGFTGSDVIRETNTGGFGAIRSVTGPILTOR 582
Qy 581 YRARIRVASTTNLRPLVTISGTRIYSINVNKTWKNKGGDLTFNTFDLATIGTATFTESNYS 640
Db 583 YRIRFRVASTIDFPFVTRGGTTINFRFRTRMWRGQESYESTVETFTFPNFTQSOD 642
Qy 641 SLTVGADSPAGSGEVYVDKFEPLPVNATPEABEDLOVAKAV-NGLFTSKKDALQTSVTD 699
Db 643 IIRTSIQLSGSEKGVLDRIEILPVNPAREABEDLEAKKAAQNLFRTRDGLQVNVTD 702
Qy 700 YQVQANLVCELSDELYPNEKMLMDAVKAEKLVQARNLLODTGNRRING--ENGWGTG 757
Db 703 YQVQANLVCELSDEYQGHDKKMLLEAVRAAKLSRERLLODPDFTNTINSTEENGWKA 762
Qy 758 STGIEVAGDVLKDRSLRITSAREIDTETPTVLYQOIDEILKPVTRVYKLGFGISSQ 817
Db 763 SNGVTISEGGFFKRALQASAR----ENPTIYQKVDASVLKPYTRRLOGDFVKSQ 818
Qy 818 DLBIKILIRHANOIVKNVPDNLPLVDLPVNSCGGIDRCSEQQYVDANLALENNGE-NGNM 876
Db 819 DLBIKILIRHANOIVKNVPDNLVDTSYDCSCGMRCEBQQVWNAQLETEHHHPMDCB 878
Qy 877 SSSHAPSFHDIGBIDLNNGTWVFKIPTNGYATIGNLELVESGPISGETLEBAQQ 936
Db 879 AAOCTHEPSSYINTGDLNASVDQGIWVVLKVRTTDGYATIGNLELVESGPISGESLREOR 938
Qy 937 QEQWQDKMAREKASEKAYAAKQADLRLPADYQDOKLANSQVEMSLAAQNLVQSIPI 996
Db 939 DNAKWNAELGRKRAIDRVYLAQAQNLHFLVDYQDQQLNPEIGLASINERASNLVESISG 998
Qy 997 VYNDALPEIIGMNYTSFELTNRLQQAQNLVYDNLNPAIPNGDFRNLGSDMNATSDVNVQOL 1056
Db 999 VYSDTLQIIGYIYIELSDRLQAQSYLTSERAVQNGEDFNSGLSDWTTTDAVQOD 1058
Qy 1057 SDTSVLVPMNNSQVQOFTVQPNRYRVLVARTARKEGVGDGYIIRDGANOTETLTNMC 1116
Db 1059 GNMHFLVLSHWDQVSQLRVNPKVYLVARTARKVGGDGYVTRDGAHHQETLTFNAC 1118
Qy 1117 DDDTGVLSADQTSYITKYTFPTPTEQVMDMSETEGVNIESVELVLEE 1167
Db 1119 DYDVNGTVYNDNSYITBEVVFYPTGMMVEVSESGFYIDSIEFTETQ 1169

RESULT 7

CLKA_BACTM STANDARD; PRT; 1215 AA.
AC Q45715;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pesticidal crystal protein cryIIA (Insecticidal delta-endotoxin
DE CryIIA(a) (Crysaline entomocidal protoxin) (137 kDa crystal protein).
GN CRYIIA OR CRYIIA(A) OR CRYIIA.
OS Bacillus thuringiensis (subsp. morrisoni).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1441;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=F190;
RX MEDLINE=96102856; PubMed=8586263;
RA Koo B.T., Park S.-H., Choi S.-K., Shin B.S., Kim J.I., Yu J.H.;
RT "Cloning of a novel crystal protein gene cryIIA from Bacillus
RT thuringiensis subsp. morrisoni";
RL PEMS Microbiol. Lett. 134:159-164(1995).
CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGIT
CC EPITHELIAL CELLS OF INSECTS. SELECTIVELY TOXIC TO ARTOGSEA RAPAE
CC AND NOT ACTIVE ON PLUTELLA XYLOSTELLA.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of

CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; U28801; BAB00376.1; --
DR HSPF; P02965; ICYI.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin C.
DR InterPro; IPR005639; endotoxin N.
DR InterPro; IPR008979; Gal_bind_Like.
DR Pfam; PF00555; endotoxin_1.
DR Pfam; PF03944; endotoxin_C_1.
DR Pfam; PF03945; endotoxin_N_1.
DR Toxin; Sporulation.
KW SEQUENCE 1215 AA; 137378 MW; 7CAF7A3311893D9B CRC64;
SQ

Query Match 44.9%; Score 2714; DB 1; Length 1215;
Best Local Similarity 46.3%; Pred. No. 6.4e-147; Indels 130; Gaps 23;
Matches 581; Conservative 194; Mismatches 351;
Qy 1 MSPNNQVEYELDASSSTVSNSVRYPLANDQTTTLQNNMYKYDLYRMSEGENPELFGNP 60
Db 1 MNSRKNEELIINALSIPAVSNHSAQMDLSPD-----ARIEDSLCAEG-----NNI 47
Qy 61 EFTISSVTQGTIGIVQVLTGALGVPPAGQIASYFIVGQLWPSSTVSVMEMIKQVED 120
Db 48 DPVFASTVQGTISAGRIKLVGLVPPAGQIASYFIVGQLWPSGR-DPWEIPEHVEQ 106
Qy 121 LI-DOKITDSVRKTALAGLQGLDLYVQKSLNMLNDRNDTRARSVVVTVYALDLP 179
Db 107 IVROQQTDSVRDPAIARLEGLGRYSYQOALSTLNDNRNDARSIRIRYIALDLDI 166
Qy 180 VAKIPSPAIQGVPELVSVAAQANLHLLLRDASIFGAEGFTFGEISTFYDQVTRTA 239
Db 167 TTAIPLESIRNEEVPLWVYAAQANLHLLLRDASLFGSEMGSSADVQYQEIYVE 226
Qy 240 QYSDYCVQVYNTGDKLKGNAASWLKTHOPRREMTLLVLDLVALFPYDTRTYPIETTA 299
Db 227 EYSNHCQVYNTGLNLRGTTAETWVRYNQPRDLTLGLVLDLVALFPYDTRTYPIETTA 286
Qy 300 QLTREVVYTD--IVPNRETSGGFCRRWSLNSDISFSVESAVIRSPHLPDILSEIEFTT 357
Db 287 QLTREVVYTDNGVVGAGPNN-----WPRNG-ASPSAENAIIRQPHLYDFELTNLTITR 339
Qy 358 RAGLPLNNTLEYLVYVGHGSIKYKNTNASSALERNYGTITSNKIKYIDL--ANKDIFQVRS 415
Db 340 RSQV--GTTIMNLWAGHRITFNRIQGGSTSEMYGAIT-NPVSVSDIPFVNRDVYRTVS 395
Qy 416 LGADLANYYAQVGVYVASFLLDKNTKSGSVGGFTSKPHTTQVCTQNTNTIDEIPPE 475
Db 396 LAGGLGSLGIRYGLTRVDFDMLFRN-HPDITVTGLFHPGHA--GIATQVKDSDTELPE 452
Qy 476 --NEPLSRGYSHRLSHITSYSPKSNASSPARYGNLVPFWTHRSADVNTVYSDKITQIP 533
Db 453 TTEQPNYRAFSLHLSHI-----SMGPTTQDVPVYVSWTHQSADRNTNTINSDRITQIP 504
Qy 534 VVKAHTLVSGTIVKGPFGTCGNILKRTSSGPLAYTQSVSVKSPLSQRYRARIYSTNL 593
Db 505 LVKAHTLVSGTIVKGPFGTCGDLIRRTSGCPFAFNVNLDNLNLSQRYRARIYSTNL 564
Qy 594 RLFVTISGTRIYSINVNKTWKNKGGDLTFNTFDLATIGTATFTESNYSLSLTVGADSPAGG 653
Db 565 RIYTVAGERIPAGQPKTMDAGAPLATQSPFSYATINTAFTFPPRSSSLTIGADTFSSGN 624

654 EYVYDFKELIPVATFEABEDLDVAKAVNGLFSTSKD-ALQTSVTVYQVNAANVECL 712
 625 EVYVDFELIQVATFEABEDLDVAKAVNGLFSTSKD-ALQTSVTVYQVNAANVECL 684
 713 SDELYPNEKMLDVAKEKELVQARNLQDTPGNRING--ENGWTSSTGLEVAEGDLV 770
 685 SDEFLDKKRELLBEVYKAKLSDERNLLQDPTFTSISGQDRGWIGSTGISIQGGDDIP 744
 771 KDRSLRLTSAREIDTETPTLYQIDESLLKPYTRYKLGKFGIGSSDLEIKLIRHRANQ 830
 745 KENYVRLPGT--VD-ECYPTLYQIDESQLKSYTRYQLRGYIEDSQDLLEIYLIRYNAKH 801
 831 IVKNVPLNLP-----DVLPNVSCGIDRCSCFQQYVDANLALENGENGMS-SDSHAPSP 885
 802 ETLSPVGTESWPSSSGVYPSGRCGEPRNRCAPR--IEWNPDLDCSCRYGEKCVHSHHPSL 859
 886 HIDEGETDLNENTGIVWVEKIPPTNGVATTGNLBEVGLSGETLERAOQOQOQODKM 945
 860 DIDVCGTDLNEDLGVWIFIKTDQGHAKLGNLEFIEBKPLLGKALSRVKAEEKWRDKY 919
 946 ARKRGASEKAYAAKQAIIDRLFADYDQOKLNSGVEMSDMLAAQNLVOSIPVYVNDALPEI 1005
 920 EKQLETRVYTEAKESVDALFVDSQYDKLOANTNIGIIGHGADKQVRIREPILSELPVI 979
 1006 PGMNYSFTLNRLOQANLIDYLNAPNGDFRNLSDMNATSDVNVQQLSDTSVLVLP 1065
 980 PSINAAIFEELEGGHIFRAYSIDYDARVINKGDFNGLSCVWNVKGVDPVQVNEHRSVLVS 1039
 1066 NNSQVSOQFTVQPNRYVLRVATKGVGGVYVLRDQANQTEITFN----- 1114
 1040 EWEAEVSKVRVCPDRGYLLRATYKGVGGVYVLRDQANQTEITFN----- 1099
 1115 --ICDDDTGVLSADOTS----- 1129
 1100 TVTCDNITNQSAGSDTACNSYRGVEDGYENRYEPNPSAPVNYTPYBEGYTDQGY 1159
 1130 -----YIKTVETPSTEQWIDMSSETGVNIESVELVLEE 1167
 1160 NHCVSDRGRNHTPLPAGYVTLLEYPPETEQWIEGETEGTFIGVSVLELLMEE 1215

RESULT 8
 C1BB BACTU
 ID C1BB BACTU STANDARD; PRT; 1229 AA.
 AC Q45739;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Pesticidal crystal protein cryIb (insecticidal delta-endotoxin
 DE CryIb(b)) (Crystalline entomocidal protoxin) (140 kDa crystal protein).
 GN CryIb or CryIb(b) OR CRYETS.
 OS Bacillus thuringiensis
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1428;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRL B-21110 / BG5847;
 RA Donovan W.P., Tan Y., Jany C.S., Gonzalez J.M. Jr.;
 RT "Bacillus thuringiensis cryet4 and cryet5 toxin genes and proteins
 RT toxic to lepidopteran insects.";
 RL Patent number US5322687, 21-JUN-1994.
 CC -!- FUNCTION: PROMOTES COLLOIDOSMOTIC LYSIS BY BINDING TO THE MIDGUT
 CC EPITHELIAL CELLS OF MANY LEPIDOPTERAN LARVAE.
 CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
 CC sporulation and is accumulated both as an inclusion and as part of
 CC the spore coat.
 CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
 CC terminus.
 CC -!- SIMILARITY: Belongs to the delta endotoxin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; L32020; AAA22344.1; -
 CC HSPF; P02965; ICYI.
 CC InterPro; IPR001178; Endotoxin.
 CC InterPro; IPR005638; endotoxin C.
 CC InterPro; IPR005639; endotoxin N.
 CC InterPro; IPR008979; Gal_bind_Like.
 CC Pfam; PF00555; endotoxin; 1.
 CC Pfam; PF03944; endotoxin; 1.
 CC Pfam; PF03945; endotoxin; 1.
 CC Toxin; Sporulation.
 CC KW
 CC SEQUENCE 1229 AA; 139769 MW; A4C949DB675C3269 CRC64;
 Query Match 42.2%; Score 2552; DB 1; Length 1229;
 Best Local Similarity 43.2%; Pred. No. 1.1e-137;
 Matches 548; Conservative 205; Mismatches 375; Indels 140; Gaps 22;
 QY 1 MSPNNOMEYEILDASSSTVSNSVRYPLANDQTTLQNNMYKYLRMSEGENPELFGNP 60
 DB 1 MTSRKNENELINALSTPTVSNPTQNLSPD-----ARIESLCAE-----VNNI 47
 QY 61 ETFISSVTQGTGIGVQVGLGALGVPFAGQIASYPSFIVGQLWPSSTVSVMEMHKVED 120
 DB 48 DPVFASTVQGTGIGVQVGLGALGVPFAGQIASYPSFIVGQLWPSSTVSVMEMHKVED 106
 QY 121 LIDQKITSVRKMTALAGLOGLDGVYQKSLKWLLENDRTRASVVVTVYALELDPV 180
 DB 107 LIROQVTEENTNTAARLEGLRGYRYSQQALETWLDNRNDRASRSLILERYVALELOIT 166
 QY 181 AKIPSPAISGOEVLVSYAQAAMHLHLLLDASIFGAENGFTGCEISTFYDRQVTRTAQ 240
 DB 167 TALPLFRIRIEVPLLVAYQAAMHLHLLLDASIFGAENGFTGCEISTFYDRQVTRTAQ 226
 QY 241 YSDYCVKYNNTGDKLKGNTAAASWLKYHQPREMTLLVLDLVALPNTDTTRYPIETTAQ 300
 DB 227 YSNHCQWYNTGLNLRGTNAESWLRYNQFRDLTLGLVLDLVALPNTDTTRYPIETTAQ 286
 QY 301 LTRVYVTDPIVPRNETSGCFRMSLNSDISPSEVESAVIRSPHLFDLILSEFYVTRAG 360
 DB 287 LTRVYVTDPIVPRNETSGCFRMSLNSDISPSEVESAVIRSPHLFDLILSEFYVTRAG 344
 QY 361 LPLANTBYLYWVGHISIKY----NTNASSALERNYGTITSNKIKYDLANKDIFQVRS 416
 DB 345 SRWSSTQHMVYVGHRLNFRPIGGLTNTSTOGLTNTSINPVLQF---TSRDVYRTE 401
 QY 417 GADLANYAQVYGVYVYASFTLLDKNTSGSGVGGTYSKPHTTMQVCTQNTYDIPPE- 475
 DB 402 AGTNILEFTPVNGVPMARFNPQ-NIYERGAITYSQYQGVGI--QLFDSFETLPPET 458
 QY 476 -NEPLSRGYSHRLSHITSYSKSNASSPARVGN--LPVPAWTHRSADVNTVYSDKIQ 531
 DB 459 TERPNYSYSHRLSHIGLI-----IGNTLRAPVYSHRSADRTNIGNRITQ 507
 QY 532 IPVKAHTLVSGTIVKPGFTGGNLRKTSGLAYTSVSKPSLSQRYARIRYASTT 591
 DB 508 IPLVKALNLSGVTVVGGPGTGGDILRRTWGTFGDIRLNLNINVELSQRYRIRYASTT 567
 QY 592 NLRPLVFTISGRYIYSINNVKTMNKGDDLTFTNTFDLATIGTFTFTSYNSDSLTVGADSPAS 651
 DB 568 DLQPFTRINGTIVNIGNFSRTMNRGDNLEYSFRTAGFTFPFNLNAQSTFTLGAQSP-S 626
 QY 652 GGEVYVDFELIPVATFEABEDLDVAKAVNGLFSTSKD-ALQTSVTVYQVNAANVECL 710
 DB 627 NQRYVIDRVEPFAEVTFAEYDIERAKVNAALFTSTNPRRLKTDVTDYHIDQVSNMVA 686
 QY 711 CLSDLEYPNKRLMDVAKVKAIRLVQARNLQDTPGNRING----- 751
 DB 687 CLSDLEYPNKRLMDVAKVKAIRLVQARNLQDTPGNRING----- 746


```
Db 621 GNGEYVVDKILIPVATFEAYDYLERAQAVNALFTNPNRLKTDVTDVHIDQVSNLV 680
Qy 710 ECLSDLEYPNKRMLMDAVKRAKRLVQARNLIQDTGFNRIN----- 750
Db 681 ACLSDPECLDKRELEKVKYAKRLSDERNLLQDPNFTSINKQDPPISTNEQSNFTSIHE 740
Qy 751 -GENGWTSGLIEAGVDVLPKDRSLRLTSARIDTETPTLYQOIDSLLKPTRYKL 809
Db 741 OSEHGWSGSENITIQEGNDVFKENTVTLPGT---FNECYPTLYQKIGESSELKAYTRYQL 797
Qy 810 KGFISGSDLEIKLHRANQIVKNVP--DNLLPDLV--PVNSCGSIDRCSQQVVDANL 865
Db 798 KGYIEDSQLELYLIRAKHETLDPGTSPLWSVESPIGRCEPNRCA--PHFBNP 855
Qy 866 ALENGENG-NMSSDSHAFSFHIDTGEIDLNENTGIWVVKIPTTNGYATLGNLSEVGG 924
Db 856 DLDSCRDGEKCAHSHHPSLDIDVGCTDLHENLGVWVVKIKTOGHARGNLSPIEBK 915
Qy 925 PLSGETLERAQOQDQWQDQWARKGSEKAYAAQALDRLPADYQDQKLSGVEMSDM 984
Db 916 PLLGALSRLVKRAEKKRDKREKLQLETKRVVTEAKEAVDALFVDSQYDLQADTNIGM 975
Qy 985 LAAQNLVQSIPIVYNDALPEIPGMVYTSFTELTNRLQAAWNLVDLRNAIPNGDFRGLSD 1044
Db 976 HAADKLVRIRAYLSELVPIPGVNAIPFEELEGGHIIITAIISLDARNVWVGDFNGLTC 1035
Qy 1045 WNAISDVANVQLSDTSVLVPIPNWSQVQFTVPQPNRYRLVLTATARKGVDGYYIIRDG 1104
Db 1036 WNVKGVHDVQSHRSDLVIPWMAEVSQAVRVCPCGGYLLAVTAYKEGVEGCVTIHEI 1095
Qy 1105 ANQETTLTP----- 1113
Db 1096 ENNTDELKPKNREBEVYPTGTCTNDYTAHQGTAGCADACNSRNAGYEDAYEVDVTTASV 1155
Qy 1114 -----NICDDDTGVLSAD--QTSVITKVEFTSTROVWIDMSETGV 1154
Db 1156 NYKPYEEETTVDRDHNCSYDRGVNPPVPVAGVYTKELSYFPETDVTWIEIGETEK 1215
Qy 1155 FNIESVELVLEE 1167
Db 1216 FIVDSVELLMEE 1228
[1]
[1]
RESULT 10
C1BC BACTM STANDARD; PRT; 1233 AA.
AC Q45774:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Pesticidal crystal protein cryIbC (Insecticidal delta-endotoxin
GN CryIb(c) (crystalline entomocidal protoxin) (140 kDa crystal protein).
OS Bacillus thuringiensis (subsp. morrisoni).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=1441;
[1]
[1]
SEQUENCE FROM N.A.
RA Bishop A.H., Bone E.J., Ellar D.J.;
RT "Cloning of novel Bacillus thuringiensis delta-endotoxin.";
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF INSECTS.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z46442; CAA86568.1; -
DR HSP; P02965; ICY.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal bind_like.
DR Pfam; PF00555; endotoxin_1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
KW Toxin; Sporulation.
SQ SEQUENCE 1233 AA; 140451 MW; 7318382413529F21 CRC64;
Query Match 41.2%; Score 2491; DB 1; Length 1233;
Best Local Similarity 42.1%; Pred. No. 3.4e-134;
Matches 536; Conservative 208; Mismatches 384; Indels 144; Gaps 20;
Qy 1 MSPNQNEYEILDASSSTSVSDNSVRYPLANDQTTTLQNMNRYKDYLRMSEGENPELFGNP 60
Db 1 MTSNRKNEEINALSIPVSNPSTQMNLSPD-----ARIEDSLCVAE-----VNNI 47
Qy 61 ETPISSTVQTGIGIGVQVGLGALGVPPAGQIASFYSFIVGQLWPSSTVSVWEMIKQVED 120
Db 48 DPFVSASTVQTGINTAGRIILGVLPVPPAGQLASFYSFLVGLWPSGR-DPEWIFLSEHVEQ 106
Qy 121 LIDQKITDSVRTALAGLQGLDGVYQKSLKNLENRNDTRASVWVVOYVIALELDV 180
Db 107 LIROQVTTENTRTAFLRLEGLRGVRSYQQALETWLDNNDARSIIILERYVALELDIT 166
Qy 181 AKIPSAISGQEVPLSVYAQAANLHLLLRDASIFGAEWGFTPGBISTFYDRQVTRTAQ 240
Db 167 TAIPFLIRNBEVPLLMVYAQAANLHLLLRDASLFGSEWGMASDVQVYQSIIRYTEE 226
Qy 241 YSDYCVKQWNTGLDKLGTNAASWLKYHQPREMTLLVLDLVALFPNYTRTYPIETTAQ 300
Db 227 YSNHCQWYNTGLNLRGTNAESWLRYNQPRDLTLGLVLDLVALFPSTYTRTYPIETSAQ 286
Qy 301 LTRVYVTDPIVFNRETSGPCRRWSLNSDISFSEVESAVIRSPHPLDILSEIEFYTRAG 360
Db 287 LTRBIYTDPIGRNAPSGFASNTWNNAPSFALERAIFRPHLLDPFEQLTIYS--AS 344
Qy 361 LPLANTEYLVWVGHISIKYK-----NTNASSALERNYGTITSNKIKYDILANKDIFQVRL 416
Db 345 SWSSTQHMVWVGHRLNFRPIGGTLNTSTQGLTNTSINPVTLPQ---TSRDVYRTESN 401
Qy 417 GADLANYAQQVGVYPYASFTLLDKNTGSGSVGGFTYKPHHTNQVCTQNTYIDEIPPE- 475
Db 402 ACTNILFTTPVNGVPMWAFNFINPQ-NIYERGATTSQPYQGVGI--QLFDSFETELPET 458
Qy 476 -NEPLSRGYSRHLSHITSYSPSKNASSPARVGN---LPVFAWTHRSADVNTNTVYSKTIQ 531
Db 459 TERPNYESYRHLSHIGLI-----IGNTLRAPVYSWTHRSADRTNIGPNRIYQ 507
Qy 532 IPVVAHRLVSGTIVIKGPGFTGGNLIKRTSSGLAYTSVSVKSPISQRYRARIYASTT 591
Db 508 IPLVKALNLHSGVTVVGGPGFTGGDILRRNTNGTGDIRLNININPLSQRYVRIRYASTT 567
Qy 592 NURLFVTTISGTRIYSINVKNTKNGKDDLTFNTFDLATIGTATFSPYSLSLTVGADSPAS 651
Db 568 DLQFFTRINGTTVNTGNFSRTNMRGDNLEYRSFRTAGFTFPNPLAQSTFTLGAQSP-S 626
Qy 652 GGEVYVDVDFELIPVNAITFEABEDLVAKVANGLTSSKKD-ALQTSVTDYQVNAANLVE 710
Db 627 NOEYIDRVFVPAVTFEAYDLERAQAVNALFTNPNRLKTDVTDVHIDQVSNMVA 686
Qy 711 CLSDLEYPNKRMLMDAVKRAKRLVQARNLIQDTGFNRIN----- 751
Db 687 CLSDDFCLDKRELEKVKYAKRLSDERNLLQDPNFTFISQQLSPASIDQSNFNPINEL 746
```



```
Qy 815 SSODLEIKLIRHRANIVKVP--DNLLPDVL--PVNSCGGIDRCSEQOYDANLALENN 870
Db 806 DSODLEYILIRYNKAKHETLDVGTGSEVWPLSVESPIGRGCEPNRCA--PHFEMNPDLDCS 863
Qy 871 GENG-NMSSDSHAFSEHIDTCEIDLENENTGTVVFKIPTNGYATLGNLELVESGPLSGE 929
Db 864 CRDEKCAHSHHFSLDIDVGCIDHENLGVVVFVKIKTQSGHARLGNLEFIEKPLLGE 923
Qy 930 TLRAQOQOQWQDMARKGASKAYAAKQADIRLPADYQDQKLSNGVEMSDMLAAQN 989
Db 924 ALSRVKRAEKWRDKREKLEKTRKRVYTEAKEAVDALPVDSDYDLQADTNGIGHAADK 983
Qy 990 LVQSPVYNDALPEIKENYVTSFELTNRLQQAANLYDLRNPNGDFRGLSDMNAIS 1049
Db 984 LVHRIRAYLSVIFGVNAEIEFELEGRIITAIISLYDAENVYNGDFNGLGACWYKG 1043
Qy 1050 DVNQVQLSDTSVLVFNWNSQVQFTQPNRYRVLRTARKEGVGQGYVIIRDCANQTE 1109
Db 1044 HVDVQOQSHRSVLVPEWAEVSVQVRVCPGRGVLIRVATYKEGVGBCVTHIEENNTD 1103
Qy 1110 TLTFNICDD-----DTGVLS-----ADQTS----- 1129
Db 1104 ELKFNCEEBEYVDTGTCDNYTAHQGTAAACNSRNAGYEDAYEVDTTASVNYKPTVEER 1163
Qy 1130 -----YITKVEETPTSEQVWIDMSFEGVNIIESVELV 1163
Db 1164 TYDVRDNHCEYDRGVYVPPVAGYMTKSELYFPFETDKYVIBIGETEGKPIVDSVELL 1223
Qy 1164 LEEB 1167
Db 1224 LMEB 1227
```

RESULT 12

```
Qy 77AA BACTU
Db 003749; STANDARD; PRT; 1138 AA.
AC 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Pesticidal crystal protein cry7Aa (insecticidal delta-endotoxin
DE CryVIIA(a)) (Crystalline entomocidal protoxin) (129 kDa crystal
DE protein).
GN CRY7AA OR CRYVIIA(A) OR CRYVIIIC.
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92384571; PubMed=1514800;
RA Lambert B., Hofte H., Annys K., Jansens S., Soetaert P., Peferoen M.;
RT "Novel Bacillus thuringiensis insecticidal crystal protein with a
RT silent activity against coleopteran larvae."
RL Appl. Environ. Microbiol. 58:2536-2542(1992).
CC -!- FUNCTION. PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF COLEOPTERA. THIS PROTEIN IS NOT TOXIC IN ITS
CC NATURAL FORM. IT IS HIGHLY TOXIC TO COLORADO POTATO BEETLE LARVAE
CC AFTER AN IN VITRO SOLUBILIZATION AND TRYPSIN ACTIVATION STEP.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch)
CC
```

```
DR EMBL; M64478; AAA22351.1; -.
DR EMBL; A07236; CAA00646.1; -.
DR PIR; A48944; A48944.
DR HSSP; P07130; 1DLC.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin C.
DR InterPro; IPR005639; endotoxin N.
DR InterPro; IPR008979; Gal_bind_like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin C; 1.
DR Pfam; PF03945; endotoxin N; 1.
DR Toxin; Sporulation.
DR KX
SQ SEQUENCE 1138 AA; 129391 MW; 69D8676D4F6AIPAC CRC64;

Query Match 40.7%; Score 2458.5; DB 1; Length 1138;
Best Local Similarity 43.5%; Pred. No. 2.2e-132;
Matches 514; Conservative 208; Mismatches 400; Indels 59; Gaps 17;

Qy 1 MSPNQNEYRILDASSSTVSVDNSVRYPLANDQTTLQNNMYKDYLRMSEGENPELPGNP 60
Db 1 MNLNLDGYE-----DSNRTLNLSINTPTQKALSPSLKNMYQDPLSITERQEPALASG 55
Qy 61 ETFISSSTVQTGIGVQVLGALGVFPAGQIASFYFIVGQLWFSSTVSVMHMKQVED 120
Db 56 NT-----AJNTVSVTGTALSALGVPCASPTNFYVKIAGLLWPENG-KIWDEPMTVEBA 109
Qy 121 LIDQKITDSVRKTMALAGLQGLDGLDYQKSLRWLENRNDTRARSVVVTVYIALELQV 180
Db 110 LIDQKIEEYVRNKAIALDLGLGSALDKYQKALADMLGKQDDPEALISVATEFRIIDLPE 169
Qy 181 AKIPSAISGOEVELLSVYQAANLHLLLDASIFGAEMGFTGCEISTFYDQVTRTAQ 240
Db 170 FSPMSFKVTGYEIPLLTVYQAANLHLLLDSTLYGDKWGTQNNIENTNRQOKR1SE 229
Qy 241 YSDYCVRYNTGLDKLGTNAASWLKYHQPREMTLLVLDLVALPFPNYDTTPTIETTAQ 300
Db 230 YSDHCTRWNSGLSRLNGSTVBQMINYRPREMLMALDLVAVPPHDPYRYSMETSTQ 289
Qy 301 LTRVYVTDPIVFNRETSGGFCRWSLNSDI--SFSEVESAVIRSPHLFDILSEIEFYTR 358
Db 290 LTRVYVTDPIVFNRETSGGFCRWSLNSDI--SFSEVESAVIRSPHLFDILSEIEFYTR 338
Qy 359 AGLPLNTE-VLEYVGVHSIKYKNTNASSALERNVGTITSNKKIK--VYDLANKDIFQVRS 415
Db 339 YKAFSHIQLPDLFWASHAKVSFKSESNLYTTGIYKTSYISGASVFGNDIYRTLA 398
Qy 416 LGADLANYAQQVGYVPYASFTLLDKNTGSGVGGFTYSKPHTTMVOCTQNTYNTIDEIPE 475
Db 399 APSVVVYPTQNYGVEQVEFY-----GVKGHVHYRGDNKYDL---TYDSIDQLPDD 446
Qy 476 NEPLSRGYSHRLSHITSYSPSKNASSPARYGN--LPVPANTHRSADVTNTVYSDKITQIP 533
Db 447 GEPIHEKYTHRLCHATAI-----FKSTPDYDNATIPISVTHRSAEYNNRYPNKIKTIP 501
Qy 534 VVKAHTLVSGTTVIKPGFTGNILKRTSSGLPAYSVSXKSPLSQRYRARIYASTNL 593
Db 502 AVRMKLDPSVTVKVGPGFTGDLVKRGSTYIGDIKATVNSPLSQKRVVRVRYATNVSG 561
Qy 594 RLFTYISG-----TRIYSINVNKNMKGDDLTFTNFDLATITGAPTFPNYSDSLTVGADS 648
Db 562 QPNVYINDKITLQTKPO--NTVETIGEGKLTYSFGYIEYSTTITQFPDDEPKITLHLS 619
Qy 649 PASGEVYVDKFLIPYNATPEAEEDLDVAKKAVNGLPTSKKDALQTSVTVYVNOAANL 708
Db 620 LSNNSYFVDSIEPIVDVYVAEKEKLEKAQAVNTLTGRRNALQKDVDTKVDQVAIL 679
Qy 709 VECLSDLEYPNKEMLMDAVKEARLVQARNLLOUTGFNRING--ENGWTSSTGIEVAG 766
Db 680 VDCISGLDLYPNKELQNLVYAKRLSVSRNLLDPTFDSINSSENGCMYSGNGVING 739
Qy 767 DVLPKDRSLRSLTSARBITDTETPTLYYQOIDESELLKPVTRYKLGKFGISSODLEIKLRH 826
Db 740 DFVFKGNVLIIFSGTN--DTQ-YPTVLYQKIDESKLEKRYKLGKFGISSODLEAVYRY 796
```


Db 909 TKRNKLAQMTTETQAIYTRAKQALDNLFAAQSHLKIQVTFABIAARKIVOSIREVY 968
 Qy 999 NDALPEIPGMYNTSFTELNRLOQANWLYDLRNAIPNGDFRGLSDMNATSDVNVQQLSD 1058
 Db 969 MSMLSVGVNHPITFELSGRQVAFQLYDYVRNVNRGRFLNGLSDMIVTSDVNVQBEAG 1028
 Qy 1059 TSVLYVFNNSQVSOFTQVQPNRYRVLVATKRGVGVYIIRGANGQTETLTFNICDD 1118
 Db 1029 NNVLVLNMDAQVLNRVKLYQDGRVLRVATARKIGIGEGYITITDEEGHTDQAFAC-E 1087
 Qy 1119 DTGVLSDAQTSYITKTVEFTSTEQWIDMETEGVFNESVLEVLLEE 1167
 Db 1088 EIDASNAFISGYITKELEFPDTEKVIHEIGETGIFLVESIEFLAEE 1136
 RESULT 14
 C1BD_BACTZ STANDARD; PRT; 1231 AA.
 AC Q92AZ5; DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pesticidal crystal protein cryIbD (insecticidal delta-endotoxin
 DE CryIb(d)) (Crystalline entomocidal protoxin) (140 kDa crystal protein).
 GN CRYIbD OR CRYIbD OR CRYIa1 OR CRYIa2.
 OS Bacillus thuringiensis (subsp. wuhanensis).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=52024;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=HD-525;
 RX MEDLINE=20153386; PubMed=10688690;
 RA Kuo W.-S., Lin J.-H., Tzeng C.-C., Kao S.-S., Chak K.-F.;
 RT "Cloning of two new cry genes from Bacillus thuringiensis subsp.
 RT wuhanensis strain";
 RL Curr. Microbiol. 40:227-232 (2000).
 CC -!- FUNCTION: PROMOTES COLLOIDOSMOTIC LYSIS BY BINDING TO THE MIDGUT
 CC EPITHELIAL CELLS OF LEPIDOPTERAN LARVAE. TOXIC TO PLOTELLA
 CC XYLOSTELLA.
 CC -!- DEVELOPMENTAL STAGES: The crystal protein is produced during
 CC sporulation and is accumulated both as an inclusion and as part of
 CC the spore coat.
 CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
 CC terminus.
 CC -!- SIMILARITY: Belongs to the delta endotoxin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U70726; AAD10292.1; -.
 DR HSSP; P02965; ICII.
 DR InterPro; IPR001178; Endotoxin.
 DR InterPro; IPR005638; endotoxin C.
 DR InterPro; IPR005639; endotoxin N.
 DR InterPro; IPR008979; Gal_bind_like.
 DR Pfam; PF00555; endotoxin; 1.
 DR Pfam; PF03944; endotoxin C; 1.
 DR Pfam; PF03945; endotoxin N; 1.
 KW Toxin; Sporulation.
 SQ SEQUENCE 1231 AA; 139654 MW; 129A0371CDDBE52 CRC64;
 Query Match 40.5%; Score 2445; DB 1; Length 1231;
 Best Local Similarity 42.0%; Pred. No. 1.4e-131;
 Matches 537; Conservative 205; Mismatches 377; Indels 160; Gaps 26;
 Qy 1 MSPNQNEYILDASSSTVSVDNSVRYPLANDQTTLQNMNMYKDYLRMSGEPBELFGNP 60
 Db 1 MTSRKNENEINALSTIPAVSNHSAQMDSLD-----ARIEDSLCIAEGNNI-----NP 49

Qy 61 BTFISSSTVQTGIGIVGVLTGALGVPPAGQIASFVSPVIGQLWSPSTVSVMIMKQVED 120
 Db 50 --LVSASTVQTGIGIAGRIILGVLPVPPAGQIASFVSPVIGQLWSPSGR-DPMEIFLEHVEQ 106
 Qy 121 LIDOKITDSVRKTAALAGIQLGDLGVYQKSLKNLENRNDTRARSVVVVVQVYALBELDFV 180
 Db 107 LIROQVTEHTANTALRLEGLGRVRSYQQALETWLDNRNDARSIIILERYVALELDIT 166
 Qy 181 AKIPSAISQGVFLLSYAAQANLHLLLRDASIFGAEWGPTPGCIISTFYDQVTRTAQ 240
 Db 167 TAIPLERIRNEEVPLLAVYAQAANLHLLLRDASIFGSEWGASDVNVQYEQIRYTEE 226
 Qy 241 YSDYCVKMYNTGLDKLKGTTNAASHKTHQPREMTLLVLDLVALPNVDTTRYPIETTAQ 300
 Db 227 YSNHCYQWYNTGLNNLRGTNAESWLRYNQFREDTLTGVLDLVALPSPYDTRYPIINTSAQ 286
 Qy 301 LTRVYVTDPIVFNRETSGGFCRRWSLNSDISFSEVSASVIRSPHLFDLSEIEFTYTRAG 360
 Db 287 LTRVIYTDPIORTNAPSGFASTNFWNNAPSAIEAAI FRPHLLDDPESQLTLYS--AS 344
 Qy 361 LPLNTEYLEYVWGHESIKYK---NTNASSALERNYGTITSNKIKYDLANKDIFQVRL 416
 Db 345 SRMSSTQHMNTWVGHRLNFRPIGGTTLNTSTQCLTNNTSINPVLQF---TSRDVYRTESN 401
 Qy 417 GADLANYAYQVYVPYASFTLLDKNTGSGVGGFTYSKPHHTMQVCTQNYNTIDILPPE- 475
 Db 402 AGTNILEFTTPVNGVPMARFNFINPQ-NIYERGATTSYQYQGVGI--QLFDSSETLPPET 458
 Qy 476 -NEPLSRGYSRHLSHITSYSFSKNASSPARYGN--LPVFAWTHRSADVNTVVSQIKITQ 531
 Db 459 TERPNYSYSRHLSHIGLI-----LGNLRAPVYSWTHRSADRTNTIGPNRIITQ 507
 Qy 532 IPVVKATHLVSGTTVIKPGPTGGNILKRTSSG-----PLATYSVSVKSPLSOR 580
 Db 508 IPAVKGRFLNG-SVISGPGTGGDVVRLNRNGNIQNRGVIEVPIQTSTST-----R 560
 Qy 581 YRAIRVASTTNLRFLTVISGTRIYSINVKTMKGGDLTFNTPLATIGTATFTSNYSD 640
 Db 561 YRVVRVYASVTSIELNVNLGNSSIFNTLPATAASLDNLQSGDGPVYVEINNAFTSATGN- 619
 Qy 641 SLTVGADSPASGGGVYVDKFLIPVNAFPAEEDLDVAKAVNGLFTSKKD-ALQTSVTD 699
 Db 620 --YVARNFSANAEEVIDRFEFIPVATFPAEYDLERAQKAVNALFTSTNPRRLATDVT 677
 Qy 700 YQVNAQANLVECLSDLEYNEKRMMLMDAVKAKRLVQARNLIQDTGFRING----- 751
 Db 678 YHIDQVSNMVACLSDLEFCLDEKRELFEKVYAKRLSDERNLIQDPNFTFISQLSFASID 737
 Qy 752 -----ENGWTSCTGIEVAGDVLFKDRSLRLTSAREIDTETVPTLYLOQIDES 799
 Db 738 GQSNFTSINELSEHGWSGSENVTIQEGNDVPEKNVYVLPQT---FNECTPNLYLOKIGES 794
 Qy 800 LLKPYTRYKLGKFGTSSQDLLEIKLIRHANOIVKNVP--DNLLPDLV--PVNSCCGIDRC 855
 Db 795 ELKAYTRYQLAGYIEDSQDLIELIYIRYNAKHETLDVPGTDSLMLPSLVKSPICRGCEPNRC 854
 Qy 856 SEQQYVDANLALNNGENG-NMSSDSHAPSFIHDTGEIDLNENTGIWVVKIPPTNGTAT 914
 Db 855 A--PHFEMNPDLDCSRDGERCAHSHHFTILDIDVGCTDLHENLGVVVPVKIKTQEGVAR 912
 Qy 915 LGNLELVEBGPISGHTLERAAQOQOQODKMKRGSEKAYAAKQOIDRLPADYQDOK 974
 Db 913 LGNLEFTEBKPLIGEARSVRAEKQWDRKKEKQLQETKRVYTEAKETVDALFVDSHYNR 972
 Qy 975 LNSGVEMSDMLAQNLYQSIPIVYNDALPEIPGMNYTSTFTELNRLOQANWLYDLRNATP 1034
 Db 973 LQADNTNIGHIADAORLVRHEAYELPELPPIPGINAVIPEELENRISTAPSLYDARNVTK 1032
 Qy 1035 NGDFRNLGLSDMNATSDVNVQQLSDTSVLVIPNNNSQVSOQFTVPQNPVYVLVATKRGV 1094
 Db 1033 NGDPNNGLSLWNVKGVHDVQQSHRSRLVPEWEAEVSOAVRVCPRGVILRVATYKEGY 1092

Search completed: June 21, 2004, 10:10:58
Job time : 27 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 10:10:21 ; Search time 90 Seconds
(without alignments)
3663.699 Million cell updates/sec

Title: US-10-089-678-1
Perfect score: 6044
Sequence: 1 MSPNNQBYEILDASSTVS.....MSETEGVNIESVELVLEE 1167

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6044	100.0	1167	5 AAU80281	AAU80281 Bacillus
2	4008	66.3	1157	2 AAR28900	Aar28900 Toxin 50C
3	4007	66.3	1157	2 AAR25997	Aar25997 Delta-end
4	4007	66.3	1157	2 AAR33768	Aar33768 Bt isolat
5	4007	66.3	1157	2 AAU06418	AAU06418 Antiscara
6	3995	66.1	1157	2 AAR27343	Aar27343 B.thuring
7	3994	66.1	1157	2 AAR44208	Aar44208 Bacillus
8	3662.5	60.6	1149	2 AAR32354	Aar32354 Coleopter
9	3662.5	60.6	1149	2 AAR51692	Aar51692 B.thuring
10	3263.5	54.0	1210	5 AAU99256	AAU99256 Bacillus
11	3246.5	53.7	1206	5 AAU99255	AAU99255 Bacillus
12	3152	52.2	1169	2 AAU06417	AAU06417 Antiscara
13	3128	51.8	1157	2 AAU84581	AAU84581 Amino aci
14	3128	51.8	1157	2 AAU84584	AAU84584 Amino aci
15	3126	51.7	1157	2 AAU84571	AAU84571 Amino aci
16	3125	51.7	1157	2 AAU84591	AAU84591 Amino aci
17	3125	51.7	1157	2 AAU84575	AAU84575 Amino aci
18	3124	51.7	1157	2 AAU84586	AAU84586 Amino aci
19	3124	51.7	1157	2 AAU84582	AAU84582 Amino aci
20	3124	51.7	1157	2 AAU84593	AAU84593 Amino aci
21	3124	51.7	1157	2 AAU84588	AAU84588 Amino aci
22	3124	51.7	1157	2 AAU84578	AAU84578 Amino aci
23	3124	51.7	1157	2 AAU84576	AAU84576 Amino aci
24	3123	51.7	1157	2 AAR48678	AAR48678 Insectici
25	3123	51.7	1157	2 AAU84570	AAU84570 Amino aci

26	3123	51.7	1157	2 AAU84566	AAU84566 Amino aci
27	3123	51.7	1157	2 AAU84569	AAU84569 Amino aci
28	3123	51.7	1157	2 AAU84568	AAU84568 Amino aci
29	3123	51.7	1157	4 AAB82178	AAB82178 Cry9C #1
30	3122	51.7	1157	2 AAU84583	AAU84583 Amino aci
31	3122	51.7	1157	2 AAU84592	AAU84592 Amino aci
32	3121	51.6	1157	2 AAU84572	AAU84572 Amino aci
33	3121	51.6	1157	2 AAU84577	AAU84577 Amino aci
34	3120	51.6	1157	2 AAU84590	AAU84590 Amino aci
35	3120	51.6	1157	2 AAU84579	AAU84579 Amino aci
36	3119	51.6	1157	2 AAU84589	AAU84589 Amino aci
37	3118	51.6	1157	2 AAU84587	AAU84587 Amino aci
38	3117	51.6	1157	2 AAU84573	AAU84573 Amino aci
39	3115	51.5	1157	2 AAU84567	AAU84567 Amino aci
40	3115	51.5	1157	2 AAU84585	AAU84585 Amino aci
41	3114	51.5	1157	2 AAU84580	AAU84580 Amino aci
42	3109	51.4	1157	2 AAU84574	AAU84574 Amino aci
43	3100.5	51.3	1156	2 AAU46857	AAU46857 Bacillus
44	3100.5	51.3	1156	2 AAU24960	AAU24960 Bacillus
45	3100.5	51.3	1156	4 AAU02034	AAU02034 B. thurin

ALIGNMENTS

RESULT 1
AAU80281
ID AAU80281 standard; protein; 1167 AA.
XX
AC AAU80281;
XX
DT 30-JUL-2002 (first entry)
XX
DB Bacillus thuringiensis insecticidal protein.
XX
KW Insecticide; transgenic; Coleoptera larvae.
XX
OS Bacillus thuringiensis.
XX
PN JP2002045186-A.
PD 12-FEB-2002.
XX
PF 03-AUG-2000; 2000JP-00236140.
PR 03-AUG-2000; 2000JP-00236140.
XX
PA (SDSB-) SDS BIOTECH CORP.
XX
DR WPI; 2002-356468/39.
XX
PT A protein having insecticidal activity, a DNA encoding said protein, and an agent and a method for preventing harmful organisms.
XX
PS Claim 1; Page 9-12; 19pp; Japanese.

XX
CC This invention relates to a crystalline protein comprising a fully defined sequence and the nucleotide sequence encoding this protein. The protein of the invention is an agent for preventing harmful organisms comprising Bacillus thuringiensis serovar galleriae SDS502, its mutant or a microbe transformed by a DNA encoding the protein. This microbe can be used to produce a protein containing the protein, or containing a protein having insecticidal activity produced by the SDS502, its mutant or a transformed microbe, a microbe which is transformed by using the above DNA and produces the above protein having insecticidal activity, a plant or a seed transformed by using the above DNA, and Bacillus thuringiensis serovar galleriae SDS502 producing a protein comprising and producing a protein showing insecticidal activity. The protein of the invention may have insecticidal activity. The agent is used for preventing Coleoptera larvae. This sequence represents the Bacillus thuringiensis insecticide protein of the invention
XX
SQ Sequence 1167 AA;

Query Match		100.0%; Score 6044; DB 5; Length 1167;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 1167; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	MSPNNQNEVEILDASSSTSVSDNSVRYPLANDQTTLLQNNYKDYLRMSGEGNPELFGNP	60
DB	1	MSPNNQNEVEILDASSSTSVSDNSVRYPLANDQTTLLQNNYKDYLRMSGEGNPELFGNP	60
QY	61	ETPFISSSTVQTGIGIVGVGLGALGVPVPAQGLASFSYFVIGQLWPSSTVSVMEMIMKQVED	120
DB	61	ETPFISSSTVQTGIGIVGVGLGALGVPVPAQGLASFSYFVIGQLWPSSTVSVMEMIMKQVED	120
QY	121	LIDQKITDSVRKTALAGLQGLDVTYQKSLKWLLENRNDTRASSVVTQVIALLELDPV	180
DB	121	LIDQKITDSVRKTALAGLQGLDVTYQKSLKWLLENRNDTRASSVVTQVIALLELDPV	180
QY	181	AKIPSPALSGQVPLLSVYQAANLHLLLRDASIFGAEMGPTGCEISTFYDQVTRTAQ	240
DB	181	AKIPSPALSGQVPLLSVYQAANLHLLLRDASIFGAEMGPTGCEISTFYDQVTRTAQ	240
QY	241	YSDYCVKMYNTGLDKLKTNAASMLKHQFREMLLLVLDLVALFPNYDTRTYPIETTAQ	300
DB	241	YSDYCVKMYNTGLDKLKTNAASMLKHQFREMLLLVLDLVALFPNYDTRTYPIETTAQ	300
QY	301	LTRVYVTDPIVFNRTSGGFCRRWSLNSDISPSEVESAVIRSPHLPDILSBIEFTYTRAG	360
DB	301	LTRVYVTDPIVFNRTSGGFCRRWSLNSDISPSEVESAVIRSPHLPDILSBIEFTYTRAG	360
QY	361	LPLNNTVEYVWGHISIKYKNTNASSALERNYGTITSNKIKYDOLANKDIPFVRSIGADL	420
DB	361	LPLNNTVEYVWGHISIKYKNTNASSALERNYGTITSNKIKYDOLANKDIPFVRSIGADL	420
QY	421	ANYAQAUVGVYASFTLLDKNTGSGSVGGFTYSKPHITTMQVCTQNYNTIDIEIPPENEPLS	480
DB	421	ANYAQAUVGVYASFTLLDKNTGSGSVGGFTYSKPHITTMQVCTQNYNTIDIEIPPENEPLS	480
QY	481	RGYSHRLSHITSYSPSKNASSPARYGNLPVPAWTHRSADVNTVYSDKLTQIPVYKAHTL	540
DB	481	RGYSHRLSHITSYSPSKNASSPARYGNLPVPAWTHRSADVNTVYSDKLTQIPVYKAHTL	540
QY	541	VSGTIVIGKPGFTGGNLRKTSGLPLAYTSVSKPLSORYRARIYASTTNLRLPVITS	600
DB	541	VSGTIVIGKPGFTGGNLRKTSGLPLAYTSVSKPLSORYRARIYASTTNLRLPVITS	600
QY	601	GTRIYSINVNKTNKGDDLTFTNTFDLATIGTAFSTSNYSDSLTVGADSPASGGYVVDKF	660
DB	601	GTRIYSINVNKTNKGDDLTFTNTFDLATIGTAFSTSNYSDSLTVGADSPASGGYVVDKF	660
QY	661	ELIPVATPEABEDLDVAKAVNGLFTSKDALQTSVTDYQVQNAANLVECLSDLYPNE	720
DB	661	ELIPVATPEABEDLDVAKAVNGLFTSKDALQTSVTDYQVQNAANLVECLSDLYPNE	720
QY	721	KEMLDVAKEARLQVARNLQDTGPNR INGENGTGSGIERSVAGDVLFKDRSLRLTSA	780
DB	721	KEMLDVAKEARLQVARNLQDTGPNR INGENGTGSGIERSVAGDVLFKDRSLRLTSA	780
QY	781	REIDTETPTLYLQIDSLKPYTRYKLKGFISGQDLLEIKLIRHRANQIVKNVPDNL	840
DB	781	REIDTETPTLYLQIDSLKPYTRYKLKGFISGQDLLEIKLIRHRANQIVKNVPDNL	840
QY	841	PVLPVNSCGGIDRCSQQYVDANLALENNNGENGNMSSDHSAPSFHIDTGEIDLNENTGI	900
DB	841	PVLPVNSCGGIDRCSQQYVDANLALENNNGENGNMSSDHSAPSFHIDTGEIDLNENTGI	900
QY	901	VWVFKIPTTNGYATLGNLELVEGPLSGTETLRAQQEQOQWQDKWARKKGASEKAYYAAK	960
DB	901	VWVFKIPTTNGYATLGNLELVEGPLSGTETLRAQQEQOQWQDKWARKKGASEKAYYAAK	960
QY	961	QAIDRLFADYQDQKLSNGVMSDMLAAQNLVQSIPIVYNDALPEIPGMNNTSFTLTNRL	1020
DB	961	QAIDRLFADYQDQKLSNGVMSDMLAAQNLVQSIPIVYNDALPEIPGMNNTSFTLTNRL	1020
QY	1021	QOANWLYDLRNPNGDFRNLGSLDNWNTSDVNVQQLSDTSLVLPNNANSQVSOQFTVQPN	1080
DB	1021	QOANWLYDLRNPNGDFRNLGSLDNWNTSDVNVQQLSDTSLVLPNNANSQVSOQFTVQPN	1080
QY	1081	YRYLVRTARKEGVGDGVYIIRDCANQOTETLTFFNICDDDTGVLSDADQTSYIITKTVEFTPS	1140
DB	1081	YRYLVRTARKEGVGDGVYIIRDCANQOTETLTFFNICDDDTGVLSDADQTSYIITKTVEFTPS	1140
QY	1141	TEQWIDMSSETGVNIESVELVLREE	1167
DB	1141	TEQWIDMSSETGVNIESVELVLREE	1167
RESULT 2			
AAR28900 standard; protein; 1157 AA.			
XX	AC	AAR28900;	
XX	DT	24-OCT-2003 (revised)	
DT	25-MAR-2003 (revised)		
DT	01-APR-1993 (first entry)		
XX	DE	Toxin 50C.	
XX	KW	Endotoxin; acarides; pest; Two Spotted Spider; mite; phytophagus.	
XX	OS	Bacillus thuringiensis; kumamotoensis PS50C.	
XX	PN	W09219106-A1.	
XX	PD	12-NOV-1992.	
XX	PF	30-APR-1992; 92MO-US003546.	
XX	PR	30-APR-1991; 91US-00693210.	
PR	13-SEP-1991; 91US-00759248.		
PR	30-SEP-1991; 91US-00768141.		
XX	PA	(MYCO) MYCOGEN CORP.	
XX	PI	Payne JW, Cannon RJC, Bagley AL;	
XX	DR	WPI: 1992-398411/48.	
DR	N-PSDB; AAQ30821.		
XX	PT	New Bacillus thuringiensis isolates and toxins - used for controlling acarid pests of livestock, fowl, stored prods. and plants.	
XX	PS	Claim 16; Page 53 + 49-51; 62pp; English.	
XX	CC	Gene sequences encoding a toxin which is active against acarides and is obtainable from B. thuringiensis isolates PS17a, PS17b, 33f2, PS2A1, PS69D1, PS86A1 and PS50C are given in AAQ30803-07 and AAQ30820-21 respectively. The toxin is a delta-endotoxin active against acarid pests, including the Two Spotted Spider mite. The isolates can be used against non-phytophagus mites such as acarid pests of livestock, fowl and stored prods. The genes can be cloned and used to transform other hosts, which can be used to control mites, or in the case of transgenic plants, be resistant to mites. (Updated on 25-MAR-2003 to correct PN field.)	
XX	CC	(Updated on 24-OCT-2003 to standardise OS field)	
XX	SQ	Sequence 1157 AA;	
Query Match 66.3%; Score 4008; DB 2; Length 1157;			
Best Local Similarity 66.1%; Pred. No. 1.7e-255;			
Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;			
QY	1	MSPNNQNEVEILDASSSTSVSDNSVRYPLANDQTTLLQNNYKDYLRMSGEGNPELFGNP	60
DB	1	MSPNNQNEVEILDASSSTSVSDNSVRYPLANDQTTLLQNNYKDYLRMSGEGNPELFGNP	60
QY	61	ETPFISSSTVQTGIGIVGVGLGALGVPVPAQGLASFSYFVIGQLWPSSTVSVMEMIMKQVED	120


```
QY 241 YSDYCVKYNATGDKLKGTTNAASMLKYHQFRREMTLLVLDLVALFPNYDTIRTYPIETTAQ 300
DB 241 YSDYCVKYNATGDKLKGTTNAASMLKYHQFRREMTLLVLDLVALFPNYDTIRTYPIETTAQ 300
QY 301 LTRVYVTDPIVFNRETSGPCRWMSLNSDIFSSEVASIVRSPLHDLILSEIEFYTTTAAQ 360
DB 301 LTRVYVTDPIVFNRETSGPCRWMSLNSDIFSSEVASIVRSPLHDLILSEIEFYTTTAAQ 360
QY 361 LPLNTEYLVYVWGHSHIKYKNT-NASSALERNYGTITSNKIKYVLDLANKDIFOVRSIGAD 419
DB 361 LPLNTEYLVYVWGHSHIKYKNT-NASSALERNYGTITSNKIKYVLDLANKDIFOVRSIGAD 419
QY 420 LANTYAYGVYPIASFTLLDKNTGSGVGGFTYSKPHFTTMQVCTQNTYNTIDEIPENB-P 478
DB 420 LANTYAYGVYPIASFTLLDKNTGSGVGGFTYSKPHFTTMQVCTQNTYNTIDEIPENB-P 478
QY 479 LSRGSHRLSHITSSFSKNSASSPARYGNLPUFANTHRSADVTNTVYSKDKITQIPVKAH 538
DB 479 LSRGSHRLSHITSSFSKNSASSPARYGNLPUFANTHRSADVTNTVYSKDKITQIPVKAH 538
QY 539 TLVSGTTVIKPGPTGGNILKARTSGPLAYTSVSVKSPLSQRYRARIYASTNLRLLFVT 598
DB 539 TLVSGTTVIKPGPTGGNILKARTSGPLAYTSVSVKSPLSQRYRARIYASTNLRLLFVT 598
QY 599 ISGTRIYSINVNKTMKGGDULFTNTFDLATIGTAFTFNSYSDSLTVGADSPASGEVYVD 658
DB 599 ISGTRIYSINVNKTMKGGDULFTNTFDLATIGTAFTFNSYSDSLTVGADSPASGEVYVD 658
QY 718 NEKRLMVDKBAKRLVQARMLLQDQFNRLNGENGWGTSGTIRVABGDVLPKDRSLRT 778
DB 718 NEKRLMVDKBAKRLVQARMLLQDQFNRLNGENGWGTSGTIRVABGDVLPKDRSLRT 778
QY 779 SAREIDTETPTLYLQOQIDESLLKPYTRYKLGFTGSSQDLEIKLIRHRAQIVKQVDPN 838
DB 779 SAREIDTETPTLYLQOQIDESLLKPYTRYKLGFTGSSQDLEIKLIRHRAQIVKQVDPN 838
QY 839 LLPDVLVNSCGGIDRCSEQQYVDANLALENGNGENGMSSDASHPSFIHDTGELDNENT 898
DB 839 LLPDVLVNSCGGIDRCSEQQYVDANLALENGNGENGMSSDASHPSFIHDTGELDNENT 898
QY 899 GIWVFKLPTTNGVATLGNLELVEGPLSGETLERAOQOQOQWODKARKGASKAYYA 958
DB 899 GIWVFKLPTTNGVATLGNLELVEGPLSGETLERAOQOQOQWODKARKGASKAYYA 958
QY 999 AKQADIRLPADYQOQKLSGVMSDMLAAQNLVQSIPIVYNDALPEIFGMNTYTSFTLTN 1018
DB 999 AKQADIRLPADYQOQKLSGVMSDMLAAQNLVQSIPIVYNDALPEIFGMNTYTSFTLTN 1018
QY 1019 RLQQAANLYDLRNATPNGDFRNLSDNATSDVNVQQLSDTSVLVPIPNWNSQVSQOFTVQ 1078
DB 1019 RLQQAANLYDLRNATPNGDFRNLSDNATSDVNVQQLSDTSVLVPIPNWNSQVSQOFTVQ 1078
QY 1079 PNRYVLRVATKCEGVGGYVIRGANGTETLTENICDDDTGLVLSADQTSYITKVEFT 1138
DB 1079 PNRYVLRVATKCEGVGGYVIRGANGTETLTENICDDDTGLVLSADQTSYITKVEFT 1138
QY 1139 PSTEQVIMDSGTGCVFNISVELVLEBE 1167
DB 1139 PSTEQVIMDSGTGCVFNISVELVLEBE 1167
QY 1129 PYTDQMWIEISETEGTFVIESVELIVDVE 1157
DB 1129 PYTDQMWIEISETEGTFVIESVELIVDVE 1157
```

RESULT 4

AAR33768
ID AAR33768 standard; protein; 1157 AA.

XX
AC AAR33768;

DT 25-MAR-2003 (revised)

```
DT 15-JUL-1993 (first entry)
XX Bt isolate Pa50C.
XX Lepidopteran-active; toxin; Ti; Ri; plant; cell; Bt.
XX Bacillus thuringiensis.
XX WO9304587-A1.
XX 18-MAR-1993.
XX 11-SEP-1992; 92WO-US007697.
XX 12-SEP-1991; 91US-00758020.
XX 08-SEP-1992; 92US-00941650.
XX (MYCO ) MYCOGEN CORP.
XX Uyeda KA, Bradfish GA;
XX WPI; 1993-100566/12.
XX N-PSDB; AAQ38653.
XX Controlling lepidopteran pests - using compen. of Bacillus thuringiensis
XX strains or plants or microorganisms transformed with their toxin genes.
XX Claim 12; Page 22-25; 38pp; English.
XX The sequences given in AAR33768-70 represent lepidopteran-active toxins.
XX The DNA encoding these sequences were used within a Ti or Ri plasmid, to
XX transform plant cells. Whole plants can then be regenerated from the
XX transformed cells. The toxin may also be produced by cloning Bacillus
XX thuringiensis (Bt). It may then be applied directly to the plant locus.
XX (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25-MAR-2003 to
XX correct PR field.)
XX SQ Sequence 1157 AA;
Query Match 66.3%; Score 4007; DB 2; Length 1157;
Best Local Similarity 66.1%; Pred. No. 1.9e-255;
Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;
QY 1 MSPNNQNEYILDASSSTSVSDNSVRYPLANDQTTLLQNNYKYDLKASGGENPELFGNP 60
DB 1 MSPNNQNEYILDASSSTSVSDNSVRYPLANDQTTLLQNNYKYDLKASGGENPELFGNP 60
QY 61 ETEFSSSTVQGTGIGVGOVIGALGVPPAGQIASFYSTVGOVLPSPSTSVSWEMIKQVED 120
DB 61 ETEFSSSTVQGTGIGVGOVIGALGVPPAGQIASFYSTVGOVLPSPSTSVSWEMIKQVED 120
QY 61 ETEFSSSTVQGTGIGVGOVIGALGVPPAGQIASFYSTVGOVLPSPSTSVSWEMIKQVED 120
DB 61 ETEFSSSTVQGTGIGVGOVIGALGVPPAGQIASFYSTVGOVLPSPSTSVSWEMIKQVED 120
QY 121 LIDQKIDTSVRKTALAGLQGLDGVYQKSLKNWLENRNDTRARSVVVTVIALELDPV 180
DB 121 LIDQKIDTSVRKTALAGLQGLDGVYQKSLKNWLENRNDTRARSVVVTVIALELDPV 180
QY 181 AKISPSPALSGQVPLLSVYQAQANLHLILRLDASIFGAEWGTGCEISFTFYDQVTRTAQ 240
DB 181 AKISPSPALSGQVPLLSVYQAQANLHLILRLDASIFGAEWGTGCEISFTFYDQVTRTAQ 240
QY 241 YSDYCVKYNATGDKLKGTTNAASMLKYHQFRREMTLLVLDLVALFPNYDTIRTYPIETTAQ 300
DB 241 YSDYCVKYNATGDKLKGTTNAASMLKYHQFRREMTLLVLDLVALFPNYDTIRTYPIETTAQ 300
QY 301 LTRVYVTDPIVFNRETSGPCRWMSLNSDIFSSEVASIVRSPLHDLILSEIEFYTTTAAQ 360
DB 301 LTRVYVTDPIVFNRETSGPCRWMSLNSDIFSSEVASIVRSPLHDLILSEIEFYTTTAAQ 360
QY 361 LPLNTEYLVYVWGHSHIKYKNT-NASSALERNYGTITSNKIKYVLDLANKDIFOVRSIGAD 419
DB 361 LPLNTEYLVYVWGHSHIKYKNT-NASSALERNYGTITSNKIKYVLDLANKDIFOVRSIGAD 419
QY 420 LANTYAYGVYPIASFTLLDKNTGSGVGGFTYSKPHFTTMQVCTQNTYNTIDEIPENB-P 478
DB 420 LANTYAYGVYPIASFTLLDKNTGSGVGGFTYSKPHFTTMQVCTQNTYNTIDEIPENB-P 478
```


QY 779 SAREIDTETPTLYXQIDBSLLKPYTRYKLGKPIGSSQDLLEIKLIRHRANQIVNVPDN 838
Db 775 GAREIDTETPTLYXQKVEGVLPKPYTRYRLRGFVGSQGLEIYTIHQNRIRVKNVPDD 834
QY 839 LLPDLVLPVNSCGGIDRCSEQOYVDANALENNNGNMGSSDSHAFSPHIDTGRIDLNENT 898
Db 835 LLPDVSPVNSDGSINRCSEQKYNSRL-----EGENRSGDAHEFSLPIDIGELDYENENA 888
QY 899 GIWVFKIPTNGYATIGNLELVEEGPLSGTTERAQOQOQODKMARKGASEKAYYA 958
Db 889 GIWVGFKITDPEGYATIGNLELVEEGPLSGDALERLQREEQWKIQWTRREEDTRYMT 948
QY 959 AKQAIIDLFPADYQDKLNSGVMSDMLAAQNLVQSIPIVYNDALPEIPGANYTSTBLTN 1018
Db 949 SKQAVRLYSYDQOQLNPVEITDLTAAQDLIQSIPIVYNNPFPPEIPGANYTKFTLTD 1008
QY 1019 RLQQAANLYDLRNPNGDFRNLSDWNATSDVNVQOLSDTSLVLPNNNSQVSOQFTVQ 1078
Db 1009 RLQQAANLYDORNAIPNGDFRNLSDWNATPGVEVQQINETSVLVLPNNNSQVSOQFTVQ 1068
QY 1079 PNRYVLRVTARKEGVDGYVIRIDGANOTETLTFTNICDDDTGVLSDADQTSYITKTVEFT 1138
Db 1069 PNQYVLRVTARKEGVGVYIRDXGNQSETLTFSASDYDTNGVYNDQYGIKTVTFI 1128
QY 1139 PSTEQWIDMSETGVPNIESVELVLEER 1167
Db 1129 PYTDQWMBISETEGTPYIESVELIVDVE 1157
RESULT 7
AA44208
ID AA44208 standard; protein; 1157 AA.
XX AC AA44208;
XX DT 24-OCT-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 09-DEC-1995 (first entry)
XX DE Bacillus thuringiensis crystal protein PS50C.
XX KW Crystal protein; delta endotoxin; acaricide; pesticide;
XX KW biological control agent; transgenic plant; crop improvement.
XX OS Bacillus thuringiensis serovar kumamotoensis; (isolate PS50C).
XX PN US5262158-A.
XX PD 16-NOV-1993.
XX PP 30-APR-1992; 92US-00876280.
XX PR 30-APR-1991; 91US-00693210.
XX PR 13-SEP-1991; 91US-00759248.
XX PR 30-SEP-1991; 91US-00768141.
XX (MYCO) MYCOGEN CORP.
XX PA Payne JM, Cannon RJC, Bagley AL;
XX PI WPI; 1993-377387/47.
XX DR N-PSDB; AAQ51704.
XX PT Controlling acaride pests e.g. two spotted spider mite - utilising
XX PT Bacillus thuringiensis delta-toxins cloned into microbe hosts.
XX PS Disclosure; Col 71-76; 42pp; English.
XX DNA encoding the insecticidal toxin can be cloned into baculo viruses and
XX CC transferred to other host microbes, preferably E. coli NM522(pMYC 2320)
XX CC NRRL B-18769, to control acaride pests, or to plants which become
XX CC resistant to the acaricide pests. Specifically, the two-spotted spider
XX CC mite (Tetranychus urticae) is controlled. (Updated on 25-MAR-2003 to

CC correct PF field.) (Updated on 24-OCT-2003 to standardise OS field)
XX Sequence 1157 AA;
QY Query Match 66.1%; Score 3994; DB 2; Length 1157;
Best Local Similarity 66.0%; Pred. No. 1.4e-254; Mismatches 240; Indels 7;
Matches 771; Conservative 144;
QY 1 MSPNNQNEYIELDASSSTSVSDNSVRYPLANDOTTTLQNNMYXDYLRMSEGENPELFGNP 60
Db 1 MSPNNQNEYIELDASSSTSVSDNSVRYPLANDOTTTLQNNMYXDYLRMSEGENPELFGNP 60
QY 61 ETPISSSTVGTGIVQVIGALGVPFAGQIASFYSPFVQGLWPSSTSVVEMIMQVED 120
Db 61 ETPISSSTVGTGIVQVIGALGVPFAGQIASFYSPFVQGLWPSSTSVVEMIMQVED 120
QY 121 LIQDKTDSVKRTALAGLQGLGVQKSLKNLNRNDTRARSVVVTVYALELDPV 180
Db 121 LVQDKIEKYVDKALAEKGLGNALDVIQSLDNLNRNDTRARSVVVTVYALELDPV 180
QY 181 AKIPSPAISSQORVPLLSVYQAANLHLLLRDASIFGAEMGFTFCEISFTFYDRQVTRTAQ 240
Db 181 SSIPSPAVSGHVBVLLAVYQAANLHLLLRDASIFGAEMGFTFCEISFTFYDRQVTRTAQ 240
QY 241 YSDYCVQYNTGLDKLGTNAASWLKTHQPRREMTLLVLDLVALPPNYDTRTYPIETTAQ 300
Db 241 YSDYCVQYNTGLDKLGTNAASWLKTHQPRREMTLLVLDLVALPPNYDTRTYPIETTAQ 300
QY 301 LTRDVTDPVFNRETSGGFCRRWSLNSDISFSVESAVIRSPHLPDILSIEFYVTRAG 360
Db 301 LTRDVTDPVFNRETSGGFCRRWSLNSDISFSVESAVIRSPHLPDILSIEFYVTRAG 360
QY 361 LPLNTEYLEYVWGHISIKYKNT-NASSALERNYGTITSNKIKYYDLANKQIFQVRSAGAD 419
Db 361 ITLNDAYINYWSGHTLKERTADSTVYTYANYGRITSEK-NSPALEDRDIFENSTVAN 419
QY 420 LANYAQVQVYASFTLLDKNTGSGVGFTYSKPHITMQVCTONTINTIDEIIPENB-P 478
Db 420 LANYQKAYGPGSWFHWKVR--GTSSTTAYLSKTHALQGCCTQVYESSDEIPLDRTPV 477
QY 479 LSRGYSRHLSHITSYSPSKNASPARGNLPVPAWTHRSADVTNTVYSDKITQIPVVKAH 538
Db 478 VASYSRHLSHITSYSPSKNASPARGNLPVPAWTHRSADVTNTVYSDKITQIPVVKAH 535
QY 539 TLVSGTIVIKGPGFTGKILKRTSSGPLAYTSVKSPLSQRYRARIYASTTLMRLFVT 598
Db 536 MLYLGSVVGPGFTGCDILKRTNPSILGTFVAVTVGSLSQRYRARIYASTTDPF-FTL 594
QY 599 ISGTIYISINVNTMKGDDLTFTNTPLATIGTAFPTFSYSDSLTVGADSPASGEVTV 658
Db 595 YLGDITIEKNRPNKTMONGASLTETPKFASFTIDFQPRETQDKILLSMGDFSSQGEVYID 654
QY 659 KFLIIPUNATFEAREDDLVAKKAVNGLTSSKDALQTSVTDYQVNOANLVECLSDLYP 718
Db 655 RIEFIPVDETYEARQDLBAAKAVNALFTNKGRLPGVTDYEVNOANLVECLSDLYP 714
QY 719 NEKRLMDAVKEARLVQARNLLQDTGFNRINENGWGTSGTIGIEVABGDVLFQDRLSLRT 778
Db 715 NEKRLMDAVKEARLVQARNLLQDTGFNRINENGWGTSGTIGIEVABGDVLFQDRLSLRT 774
QY 779 SAREIDTETPTLYXQIDBSLLKPYTRYKLGKPIGSSQDLLEIKLIRHRANQIVNVPDN 838
Db 775 GAREIDTETPTLYXQKVEGVLPKPYTRYRLRGFVGSQGLEIYTIHQNRIRVKNVPDD 834
QY 839 LLPDLVLPVNSCGGIDRCSEQOYVDANALENNNGNMGSSDSHAFSPHIDTGRIDLNENT 898
Db 835 LLPDVSPVNSDGSINRCSEQKYNSRL-----EGENRSGDAHEFSLPIDIGELDYENENA 888
QY 899 GIWVFKIPTNGYATIGNLELVEEGPLSGTTERAQOQOQODKMARKGASEKAYYA 958
Db 889 GIWVGFKITDPEGYATIGNLELVEEGPLSGDALERLQREEQWKIQWTRREEDTRYMT 948
QY 959 AKQAIIDLFPADYQDKLNSGVMSDMLAAQNLVQSIPIVYNDALPEIPGANYTSTBLTN 1018

```

Db 949 SKQAVRLYADYQDLQNLDPVETITDLTAQDLIQSIPTVYNEMFPEIFGMNFKTELTLD 1008
Qy 1019 RLQQAANLYDLRNLATPNCDFRNLSDMNATSDVNVQQLSDTSVLVIPNNNSQVSQFTVQ 1078
Db 1009 RLQQAANLYDQNALPNCDFRNLSDMNATSDVNVQQLSDTSVLVIPNNNSQVSQFTVQ 1068
Qy 1079 PNRYVLRVTRARKEGCVGYYIIRDGANQTELTFTNLCDDDTGVLASADQTSYITKVEFT 1138
Db 1069 PNQVLRVTRARKEGCVGYYIIRDGANQTELTFTNLCDDDTGVLASADQTSYITKVEFT 1128
Qy 1139 PSTQVWIDMSETEGVFNIESVELVLEE 1167
Db 1129 PYTQMWIEISETEGTFYIESVELVDVE 1157

```

RESULT 8

AAR32354
ID AAR32354 standard; protein; 1149 AA.

XX AAR32354;

AC 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 14-JUN-1993 (first entry)

XX Coleoptera toxin from B.t. serovar japonensis variety Buibui.

XX Delta-endotoxin; larvae; Anomala cuprea; spore; crystal; Bacillus;

KW thuringiensis; virus; insect; FERM BP-3465.

XX Bacillus thuringiensis; serovar japonensis variety Buibui.

OS W09303154-A1.

PN 18-FEB-1993.

XX 31-JUL-1992; 92WO-US006404.

XX 02-AUG-1991; 91JP-00193810.

PR 23-JUL-1992; 92US-00915203.

XX (MYCO) MYCOGEN CORP.

PA (KUBI) KUBOTA CORP.

XX Ohba M, Iwahana H, Sato R, Suzuki N, Ogiwara K, Sakanaka K;

PI Hori H, Asano S, Kawasaki T;

XX WPI; 1993-076511/09.

DR N-PSDB; AAQ36866.

XX New strain of Bacillus thuringiensis serovar japonensis - producing toxin

PT active against coleoptera larvae.

XX Claim 5; Page 28; 48pp; English.

XX The protein sequence is that of a toxin active against Coleoptera that is
CC produced from a pure culture of Bacillus thuringiensis serovar japonensis
CC variety Buibui (FERM BP-3465). The toxin is a delta-endotoxin which has a
CC mol. wt. of ca. 130 kD. It is useful for control of coleoptera larvae
CC e.g. it is effective against Anomala cuprea but has little effect on
CC Lepidoptera. The toxin can be used as B.t. spores or crystals, as opt.
CC treated cells (B.t. or transformed microorganisms) or it is expressed by
CC plants. (Updated on 25-MAR-2003 to correct FN field.) (Updated on 24-OCT-
CC 2003 to standardise OS field)

XX Sequence 1149 AA;

Query Match 60.6%; Score 3662.5; DB 2; Length 1149;

Best Local Similarity 63.8%; Pred. No. 1.1e-232;

Matches 752; Conservative 109; Mismatches 257; Indels 61; Gaps 21;

Qy 1 MSPNNQNEYEILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSGENPELFGNP 60

```

Db 1 MSPNNQNEYEILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSGENPELFGNP 60
Qy 61 BTFTSS-STVOTGIGVGVGALGVPPAGQASFPVSPVIGVQLWPSSTSVVMEIMKQVE 119
Db 61 GTFISAQAVAGTGIDIVSTIISGIGIPVLGEVFSILGSLIGLLMFSNNENNVQIPFNRVE 120
Qy 120 DLIDQKITDSVRKTALAGLQGLDGLDVYQSKLKNWLENRNDTRARSVVVTVYIALELDF 179
Db 121 ELIDQKILDSVRSRAIADLANSRIAVEVYQNALEDWRKNPHSTRSAALVKERFGNAEAIL 180
Qy 180 VAKIPSPALSCQEVPLLSVYAAQANHLHLLLDASIFGAENGFTPECEISTFVDQVTRTA 239
Db 181 RTNMGFSQNTYETPLPFTYAQAASLHLLVNRDVQIYCKENGYPQNDLDFYKEQVSYTA 240
Qy 240 QYSDYCVKYNVTGLDKLGTNAASWLYKHQFRREMTLLVLDLVALPPNYDTTYPIETTA 299
Db 241 RYSDHCQVYNAGLKLGTAKQWVDNRRFRRENWVLDLVALPPNYDARIYPLETNA 300
Qy 300 QLTREVYDPIVFNRETSGGFCRRWSLN-----SDI-----SFSVESAVIRSPHLFOIL 349
Db 301 ELTREIPTDPV-----GSVVTGQSTLISWYDMIPAALPSFSTLEN-LLKPPDFTLL 352
Qy 350 SEIEPFTTRAGLPLNNT-EYLEYVGVGHSIKYKNTWASSALERNYGTITSNKIKYYDLANK 408
Db 353 QEIRMYTS---FRQNGTIEYNYMGQRLTSLYIYGSSP--NKYSGVLAGEAEDIIIPVGQN 407
Qy 409 DIFQVRSIGADLANVYAOVGVVPASPTLLDKNTGSGVGGVPTYSKPHPTMVCQNTVYT 468
Db 408 DIYRV--VWTYIGRTNLSLLGVNVPVP-YPSNTOK-----TYSKP-----KQFAGGIKT 454
Qy 469 ID---EIPPEPEPLSRGSHLSHITSYSPKSNASSPARYGNLVPAMTHRSADVNTVY 525
Db 455 IDSGEELTYEN---YQSYSHVSVITSPEIKSTGTV--LGWVPIFGWTHSSASRNNPIY 509
Qy 526 SDKJTIQIPVTAHTLVSGTTVIKGPQ-FTGGNILKRTSSGPLAY---TSVSVKSPLSQRY 581
Db 510 ATKISQIPINKASRTSGGAVWNPQBLGNGYGPWMLSGSGSQVINLRVATDAGK-ASQRY 568
Qy 582 RARIRYASTTNLRPLPVTIS-----GTRIYSINVKTNKMGDDTLNFTDLATIGTA-P 633
Db 569 RIRIRYASDRAGKP--TISRSPEPATYSIASIATNTMTSNASLITYSTFAESGPINL 626
Qy 634 TFSNYSLSLTVGADSPASGGBVYVDKPELIPVNATFEABEDLDVAKKAVNGLFTSKOAL 693
Db 627 GISGSSRTFDISITKEAGANLYIDRIEPIVNTLFEABEDLDVAKKAVNGLFTNEKDAL 686
Qy 694 QTSVTDYQVQAANLVECLSDLELYPNEKRLMDAWEAKRLVQARNLLQDTGFNRINGEN 753
Db 687 QTSVTDYQVQAANLIECLSDLELYPNEKRLMDAWEAKRLVQARNLLQDTGFNRINGEN 746
Qy 754 GWTGSGTGEVARGDVLFKDRSLRLTSARBIDTETPTLYQQIDSLKLPYTRYKLGPI 813
Db 747 GWTGSGTGEVARGDVLFKDRSLRLTSARBIDTETPTLYQQIDSLKLPYTRYKLGPI 806
Qy 814 GSSQDLLEKLIRHRANQIVKQVNPDLPLDVLVPSNCGGIDRCSEQQYVDANLALNNGEN 873
Db 807 GSSQDLLEKLIRHRANQIVKQVNPDLPLDVLVPSNCGGIDRCSEQQYVDANLALNNGEN 866
Qy 874 GNMSSDSHAPSFHDITGIDLNENTGIIWVFKIPTTNGYATLGNLELVEEGPLSGETLER 933
Db 867 GNMSSDSHAPSFHDITGIDLNENTGIIWVFKIPTTNGYATLGNLELVEEGPLSGETLER 926
Qy 934 AQOQEQWQDQWARKRGASEKAYYAAQAIIDRLFADYQDQKLNLSGVMSDMLAAQNLVQS 993
Db 927 AQOQEQWQDQWARKRGASEKAYYAAQAIIDRLFADYQDQKLNLSGVMSDMLAAQNLVQS 986
Qy 994 IPYVYNDALPEIPGMNYTSPTELTNRLQQAANLYDLRNLATPNCDFRNLSDMNATSDVNV 1053
Db 987 IPYVYNDALPEIPGMNYTSPTELTNRLQQAANLYDLRNLATPNCDFRNLSDMNATSDVNV 1046
Qy 1054 QQLSDTSVLVTPNNNSQVSQFTVQPNRYVLRVTRARKEGCVGYYIIRDGANQTELTFT 1113

```

Db 1047 QQLSDTSVLVPMNSQVSOQFTVQPNRYRVLRTARKEGVGDGVYIIRGANQTETLTP 1106
 QY 1114 NICDDDTGVLSDQTSYITKTVFTSTPTEQVWIDMSETE 1152
 Db 1107 NICDDDTGVLSDQTSYITKTVFTSTPTEQVWIDMSETE 1145

RESULT 9
 AAR51692
 ID AAR51692 standard; protein; 1149 AA.
 AC AAR51692;
 XX
 DT 16-OCT-2003 (revised)
 DT 04-NOV-1994 (first entry)
 XX
 XX B.thuringiensis serovar Japonensis insecticidal protein.
 XX insecticidal protein; Coleoptera larvae; Buibui fungus.
 XX
 OS Bacillus thuringiensis; (serovar Japonensis.
 OS strain Buibui).
 PH Key Location/Qualifiers
 FT Region 751..766
 FT /notes="amino acids 751-766 are not given in the
 FT specification and so have been decoded from AAQ58975"
 XX
 XX JP06065292-A.
 XX 08-MAR-1994.
 XX 11-AUG-1992; 92JP-00213886.
 XX 11-AUG-1992; 92JP-00213886.
 XX (KUBI) KUBOTA CORP.
 XX WPI; 1994-121220/15.
 XX N-PSDB; AAQ58975.
 XX Insecticidal protein and DNA from Bacillus thuringiensis serovar
 PT Japonensis strain Buibui - useful in insecticides against Coleoptera
 PT insects.
 XX
 PS Claim 1; Page 9-13; 18pp; Japanese.
 XX
 CC This insecticidal protein has activity against Coleopteran insect larvae
 CC and has been isolated from Bacillus thuringiensis serovar japonensis
 CC strain Buibui. (Updated on 16-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 1149 AA;
 Query Match 60.68; Score 3662.5; DB 2; Length 1149;
 Best Local Similarity 63.84; Pred. No. 1.1e-232;
 Matches 752; Conservative 109; Mismatches 257; Indels 61; Gaps 21;

QY 1 MSPNQEYIILDASSSTVSNSRYPLANDQTTTLQNNYKDYLRMSGENPELFGNP 60
 Db 1 MSPNQEYIILDASSSTVSNSRYPLANDQTTTLQNNYKDYLRMSGENPELFGNP 60
 QY 61 ETPISS-STVQTGIGIYGOVLGALGVPPAGQIASFYFIVGQLWPSSTVSWEIMKQVE 119
 Db 61 GTFISAQDAVGTDIVSTIISGLGIPVLGEVFSILGSLGLWPSNNENWQIPNRVE 120
 QY 120 DLIDQKITDSVRKTALAGLQGLDGVYQKSLKNWLENRNDTRARSVVVTVQIAELDF 179
 Db 121 ELIDQKITDSVRKATLANSRIAYEYQNALEDWKNPHSTRSAALVKERFGNAAIL 180
 QY 180 VAKIPSAISGQEVPLLSVQAQANLHLLLRDASIFCAEWGFTPGBISTFYDQVQTRTA 239
 Db 181 RTNMGFSQTNVETPLPTTAAQASLHLLWRDVQIYKKEWGPQNDIDLFPYKEQVSYTA 240

QY 240 QYSDYCVKWNVTGLDKLKTNAASWLKCHOPREMTLLVLVLVALFPNYDTRTYPIETTA 299
 Db 241 RYSDHCYQWYNAGLNLKRGCAKQVDVFRFRMNVMVLVALFPNYDARIYPLETNA 300
 QY 300 QLTREYVTDPIVFNRETSQGFCCRNLSN-SDI-----SFSEVSESAVIRSPHLPDIL 349
 Db 301 ELTREIFDTPV-----GSVVTGQSTLSWYDMPAALPSPSTLEN-LLRKDPDFTL 352
 QY 350 SEIEFYTTAGLPLNNT-EYLEVYVGHSHIKYKTNASSALERNYGTITSNKIKYYDLANK 408
 Db 353 QEIRMYTS---PRQNGTIEYNYNGQRLTSLSYIGSSF--NKISGVLAGAEDIIIPVQGN 407
 QY 409 DIFQVRSGLADLANYYAQVYGPVYASPTLLDKNTGSCSGVGTYSKPHHTMQVCTQNTY 468
 Db 408 DIYRV--VMTYIGRYTNSLLGVNPTVF-YFSNNYTK-----TYSKP-----KQFAGGAKT 454
 QY 469 ID---EIPPENBPLSGYSHRLSHITSYFSKSNASSPARVGNLPVFAWTHRSADVNTVY 525
 Db 455 IDSGEELTYEN---YQSYSHRVSYITSFBIKSTGGTV--LGVVPIFGWTHSSASRNNFIY 509
 QY 526 SDKITQIPVYKAHTLVSGTIVIKGPG-PTGNIILKRTSSGPLAY---TSVSVKSPLSORY 581
 Db 510 ATKISQIPINKASRTSGGAVMNPQGLYNGCPVMKLSGSGSQVINLRVATDAKG-ASORY 568
 QY 582 PARIRYASTTNMLRFPVTIS-----CTRIYSINVNKTMKGGDLTPNTFDLATIGTA-F 633
 Db 569 RIRIRYASDRAGKP--TISRSPENPATYSASIAITMTSTNASLTSTPAYAESGPINL 626
 QY 634 TFSNYSDSLTVGADSPASGGEVYVDKPELIPVNAATFBAEEDLDVAKKAVNGLFTSKDAL 693
 Db 627 GISGSRTPDISITKEGAANLYIDRIEFPVNTLFEAEEDLDVAKKAVNGLFTNEKDAL 686
 QY 694 QTSVTDYQVNOAANLVECLSDELYPNKRMWDVAKKRLVQARNLLQDTGFPNRINGEN 753
 Db 687 QTSVTDYQVNOAANLVECLSDELYPNKRMWDVAKKRLVQARNLLQDTGFPNRINGEN 746
 QY 754 GWTGTCGIEVAREGDVLPKDRSLRTSAREIDTETYPYLYQOIDEKLLKPYTRYKLGKFI 813
 Db 747 GWTGTCGIEVAREGDVLPKDRSLRTSAREIDTETYPYLYQOIDEKLLKPYTRYKLGKFI 806
 QY 814 GSSODLEIKLIRHRANQIVKNVNDNLDPVLPVNSCGGIDRCSEQQYVDANLALENNEN 873
 Db 807 GSSODLEIKLIRHRANQIVKNVNDNLDPVLPVNSCGGIDRCSEQQYVDANLALENNEN 866
 QY 874 GNMSSDSHAPSHI DTGEIDLNENVTGIVVFKIPTTTNGYATLGNLELVEBGLSGETLER 933
 Db 867 GNMSSDSHAPSHI DTGEIDLNENVTGIVVFKIPTTTNGYATLGNLELVEBGLSGETLER 926
 QY 934 AQOQEQWQDMARKKASAKAYAAKQADRLPADYQDQKNSGVEMSDMLAAQNLVQS 993
 Db 927 AQOQEQWQDMARKKASAKAYAAKQADRLPADYQDQKNSGVEMSDMLAAQNLVQS 986
 QY 994 IPVYNDALPEIPGMNYTSTFELTNRLQAAWNLVDLRAIIPNGDFRNLGSDMNATSDVNV 1053
 Db 987 IPVYNDALPEIPGMNYTSTFELTNRLQAAWNLVDLRAIIPNGDFRNLGSDMNATSDVNV 1046
 QY 1054 QQLSDTSVLVPMNSQVSOQFTVQPNRYRVLRTARKEGVGDGVYIIRGANQTETLTP 1113
 Db 1047 QQLSDTSVLVPMNSQVSOQFTVQPNRYRVLRTARKEGVGDGVYIIRGANQTETLTP 1106
 QY 1114 NICDDDTGVLSDQTSYITKTVFTSTPTEQVWIDMSETE 1152
 Db 1107 NICDDDTGVLSDQTSYITKTVFTSTPTEQVWIDMSETE 1145

RESULT 10
 AAR59256
 ID AAR59256 standard; protein; 1210 AA.
 XX
 AC AAR59256;
 XX
 DT 07-OCT-2002 (first entry)
 XX


```

DE Bacillus thuringiensis Cry1218-2 protein sequence.
XX Pesticidal; spraying; dusting; broadcating; seed coating; insect pest;
KW Colorado potato beetle; western corn rootworm; southern corn rootworm;
XX insect target range; endotoxin; Cry1218.
XX Bacillus thuringiensis.
OS
XX WO200234774-A2.
FN
XX
PD 02-MAY-2002.
XX
XX 24-OCT-2001; 2001WO-US045468.
PF
XX
XX 24-OCT-2000; 2000US-0242838P.
PR
XX 23-OCT-2001; 2001US-00032717.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
PA
XX
XX Abad AR, Duck NB, Feng X, Flannagan RD, Kahn TW, Sims LE;
FI
XX
XX WPI; 2002-519178/55.
DR
XX N-PSDB; ASK87235.
DR
XX
XX New isolated pesticidal polypeptide useful for impacting insect pest e.g.
PT Colorado potato beetle.
PT
XX
XX Claim 4; Page 103-105; 176pp; English.
PS
XX
CC The present invention relates to a new pesticidal polypeptide. The
CC invention is useful for impacting an insect pest by applying the the
CC molecules of the invention to the environment of the insect pest by
CC spraying, dusting, broadcating, or seed coating, where the insect pest
CC is selected from Colorado potato beetle, western corn rootworm or
CC southern corn rootworm. The invention is also useful for increasing
CC insect target range and for producing transgenic microorganisms and
CC plants that express the pesticidal polypeptide. The invention is also
CC useful for producing transformed plants and in transforming any organism
CC to produce the pesticidal polypeptide of the invention. The present amino
CC acid sequence represents a Bacillus thuringiensis wild-type Cry1218
CC endotoxin protein
XX
XX Sequence 1210 AA;

Query Match 54.0%; Score 3263.5; DB 5; Length 1210;
Best Local Similarity 53.4%; Pred. No. 2.6e-206;
Matches 667; Conservative 178; Mismatches 284; Indels 119; Gaps 21;

QY 1 MSPNQNEVEIILDASSSTSVSDSVRYPLANDQTTTLQNMNTKYLKMSGCEGPELFGNP 60
DB 1 MSPNQNEVEIILDASSSTSVSDSVSDSVRYPLANDQTTTLQNMNTKYLKMSGCEGPELFGNP 60
QY 61 ETPFIS-SSTVQNGIGVGVGLGALGVPPAGQIASFYSPTVGLWPSSTSVVEMIMKOVE 119
DB 61 EVLVSGQDAKAAIDIVGKLLSGLGVPPGPIVSLYTLQILDLNPSGQKQWEIEMQVE 120
QY 120 DLIDQKITDSVRKTLAGLQGLGDLDVYQKSLKXWLENRNDTRARSVVVYQYIALELDP 179
DB 121 ELINQKIAEYARNKALSELGLGNQYQLYLTALKEWKEPNSRALRDVNRNFEILDLSL 180
QY 180 VAKIPSPALSGQEVPLLSVYAAQANLHLLLDASIFGAEWGFTTCEISTFYDRQVTRTA 239
DB 181 TQMFPSFRVTNPEVPLTAVYVYQAAANLHLLLDASIFGAEWGFTTINNYDQKQLTA 240
QY 240 QYSDCVKWNNTGLDKLKGTAASLKYHQFRREMTLLVLDLVALLPNNYDRTYPIETTA 299
DB 241 EYSDHCWKYETGLAKLKTSAKQWVDYQFREMVLTVLDVVALLPNNYDRTYPIETTA 300
QY 300 QLTREYVYDPIVFNRETPGCGFCRRMSLNSDISFSEVESAVIRSPHLFDILSIEFYETRA 359
DB 301 QLTREYVYDPLGAVNVSSIG--SW-YDKAPSGFVIESVIRPPHVFYITGLTVYTSQR 356
QY 360 GLPLNTEYLETWVGHISIKYKNTNASSALERNYGITTS-NKIKYVDLANKDIFQVRSIGA 418

```

```

Db 357 S--ISSARYIRHWAGHQISYHRIFSDNIIKMGYGNQNLHSTSTFDFTNYDYIKLSKDA 414
QY 419 DLAN-----YYAQVYGVFYASFLLDK--NTGSGSVGGFTYSKPHHTMQVCTQNTYNTID- 470
Db 415 VLLDIVPGYTYIPFGMEVEFFVMVQNLNTRK-----TLKNPVSVK 456
QY 471 -----BIPPE--NEPLSRGYSRHLSHITSYFSKQASPARVGNLVPFATHS 517
Db 457 DIIAGTRDSELELPPETSDQPNYESYRHLCHITSIPATGSTT-----GLVPVFSWTHKS 511
QY 518 ADVTNTVYSDKITQIPVVKAKHTL---VSG---TTVIKGPFTGGMILKRTSGPLAYTSV 571
Db 512 ADLINAVHSDKITQIPVVKVSDLAPSTGGPNNTVSGPFTGGGIIKVRNG-VIISIM 570
QY 572 SVK-SPLSQRYRARIYASTTNLRLFVITISGTRIYSINVNKTMKGGDOLTFTNTPOLATIG 630
Db 571 RVKISDINKEYSMRIRYASANTBFYINPSENVKS-HAQTKMRGEALTYNKNFYATL- 628
QY 631 TAFTPSNYSDSLTVGADSFAS---GGEVYVVKPHELIPVNATFEAREEDLDVAKAVNGLFT 687
Db 629 PPIKFTTTPFTITLGAIFEAEEDFLGIEAYIDRIEIPVDIETYEAEQDLSEAKKANALFT 688
QY 688 SKKDALQTSVTDYQVNOQANLVECLSDILYPNBKRLMDAVKEAKRLVQARNLLQDQFN 747
Db 689 NTKGLRPGVTDIEVNOQANLVECLSDILYPNBKRLLFDAREAKRLSEARNLLQDQFQ 748
QY 748 RINGENGWTSIGTIEVAREGDVLPKDRSLRLTSAREIDTETPTTYLYQQIDRSLKPYTRY 807
Db 749 EINGENGWTSIGTIEVAREGDVLPKDRSLRLTSAREIDTETPTTYLYQVBEGLKPYTRY 808
QY 808 KLKPGIGSSODLEIKLIRHRANQIVKNVDPNLLPVLVNSCGGIDRCSEQQVVDANLAL 867
Db 809 RLKRGVSSQGLEIFTHRHQTNRIVKNVDPDILLPVPVNSDGSINRCSEQKTYVNSRLEV 868
QY 868 ENNGENGWSSDSHAFSPHIDTGBEIDLNENTGIWVFKIPTTNGYATLGNLELVBEGLPS 927
Db 869 ENR-----SGBAHEFSIPIDTGBEIDYENAGIWWGFKLTDPEGATLGNLELVBEGLPS 922
QY 928 GTTLERAQQEQOQWODKMKARKGASEKAYAAKQAIIDRLFADYQDQKLSGVEMSDMLAA 987
Db 923 GDALERLQREQQWKIQMTERREEDRRYMAASKQAVDRLYADYQDQQLNPDVEITOLTA 982
QY 988 QNLVQSIPIVYNDALPEIRGMNYTSTFELTNLQQAANLYDLNAILPNGDFRNLGSLDMA 1047
Db 983 QDLIQSIPIVYVNMENFPEIRGMNYTSTFELTNLQQAANLYDLNAILPNGDFRNLGSLDMA 1042
QY 1048 TSDVNVVQQLSDTSVLVIWNNSQVSOQFTVQPNRYVRLVATRKEGVGQGYVIIRDGANG 1107
Db 1043 TPGVEVQQLNHTSVLVIWNDEQVSOQFTVQPNRYVRLVATRKEGVGQGYVIRDDGNG 1102
QY 1108 TETLTFNICDDDT-GVLSA----- 1125
Db 1103 TETLTFASDVTNGYNTQVSTNGYNTNAYNTQASSTNGYNANNMTNTQASNTNGYN 1162
QY 1126 -----DOTSYITKTVETPSTEQWIDMSSETGEVFNIESVLVEEB 1167
Db 1163 TNSVNDQGYITKTVETPSTEQWIDMSSETGEVFNIESVLVEEB 1210

```

```

RESULT 11
AAU99255
ID AAU99255 standard; protein; 1206 AA.
XX
AC AAU99255;
XX
DT 07-OCT-2002 (first entry)
XX
DE Bacillus thuringiensis Cry1218-1 protein sequence.
KW Pesticidal; spraying; dusting; broadcating; seed coating; insect pest;
KW Colorado potato beetle; western corn rootworm; southern corn rootworm;
KW insect target range; endotoxin; Cry1218.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 10:12:40 ; Search time 402 Seconds
(without alignments)
819.552 Million cell updates/sec

Title: US-10-089-678-1
Perfect score: 6044
Sequence: 1 MSPNNQNEYILDASSSTSV.....MSETGVNIESVELVLEEE 1167

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1163542 seqs, 282313646 residues

Total number of hits satisfying chosen parameters: 1163542

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6044	100.0	1167	12	US-10-089-678-1
2	3263.5	54.0	1210	13	US-10-032-717-4
3	3263.5	54.0	1210	14	US-10-414-637-4
4	3263.5	53.7	1206	15	US-10-606-320-4
5	3246.5	53.7	1206	16	US-10-032-717-2
6	3246.5	53.7	1206	17	US-10-414-637-2
7	3246.5	53.7	1206	18	US-10-606-320-2
8	3100.5	51.3	1156	14	US-10-099-285-72
9	3100.5	51.3	1156	15	US-10-428-961-28
10	2884.5	47.7	1152	15	US-10-614-524-6
11	2825.5	46.7	1150	14	US-10-099-285-74
12	2710	44.8	1134	14	US-10-099-285-76
13	2519.5	41.7	1207	15	US-09-988-462-7
14	2457	40.7	1227	15	US-10-428-961-63
15	2454.5	40.6	1186	9	US-09-826-660-23

16	2449.5	40.5	1228	15	US-10-428-961-38
17	2446.5	40.5	1228	15	US-10-614-524-2
18	2230.5	36.9	1177	9	US-09-873-873-10
19	2230.5	36.9	1177	9	US-09-873-873-12
20	2230.5	36.9	1177	9	US-09-873-873-14
21	2230.5	36.9	1177	10	US-09-997-914-10
22	2230.5	36.9	1177	10	US-09-997-914-12
23	2230.5	36.9	1177	10	US-09-997-914-14
24	2230.5	36.9	1177	14	US-10-365-645-10
25	2230.5	36.9	1177	14	US-10-365-645-12
26	2230.5	36.9	1177	14	US-10-365-645-14
27	2230.5	36.9	1177	16	US-10-672-163-10
28	2230.5	36.9	1177	16	US-10-672-163-12
29	2230.5	36.9	1177	16	US-10-672-163-14
30	2229.5	36.9	1177	9	US-09-873-873-26
31	2229.5	36.9	1177	10	US-09-997-914-26
32	2229.5	36.9	1177	14	US-10-365-645-26
33	2229.5	36.9	1177	16	US-10-672-163-26
34	2227.5	36.9	1163	9	US-09-756-526A-2
35	2227.5	36.9	1163	14	US-10-345-020-2
36	2227.5	36.9	1163	14	US-10-342-821-2
37	2225.5	36.8	1177	9	US-09-873-873-28
38	2225.5	36.8	1177	10	US-09-997-914-28
39	2225.5	36.8	1177	14	US-10-365-645-28
40	2225.5	36.8	1177	16	US-10-672-163-28
41	2215.5	36.7	1177	9	US-09-873-873-34
42	2215.5	36.7	1177	14	US-10-365-645-34
43	2191	36.3	1156	14	US-10-099-285-70
44	2154.5	35.6	1174	9	US-09-826-660-6
45	2154.5	35.6	1174	12	US-09-837-961-8

ALIGNMENTS

RESULT 1
US-10-089-678-1
; Sequence, Application US/10089678
; Publication No. US20030017967A1
; GENERAL INFORMATION:
; APPLICANT: ASANO, Shinichiro
; TITLE OF INVENTION: PROTEIN HAVING INSECTICIDAL ACTIVITY, DNA ENCODING THE PROTEIN, NOXIOUS ORGANISM-CONTROLLING AGENT AND METHOD
; FILE REFERENCE: Q68821
; CURRENT APPLICATION NUMBER: US/10/089,678
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: JP 2000-236140
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: PCT/JP01/06660
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1167
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-089-678-1

Query Match	100.0%	Score 6044	DB 12	Length 1167
Best Local Similarity	100.0%	Pred No. 0		
Matches 1167	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	MSPNNQNEYILDASSSTSVNSVRYPLANDQTTTLQNNYKDYLRMSEGENPELFGNP	60	
DB	1	MSPNNQNEYILDASSSTSVNSVRYPLANDQTTTLQNNYKDYLRMSEGENPELFGNP	60	
QY	61	ETPFISSSTVGTGIGVIGVGLGALGVPFAGIASFYSPIVQGLMPSSTVSWEMIMQVED	120	
DB	61	ETPFISSSTVGTGIGVIGVGLGALGVPFAGIASFYSPIVQGLMPSSTVSWEMIMQVED	120	
QY	121	LIDQKITTSVRKLTALAGLQGLDGVYQKSLKNWLENNDTRARSVVVTVYIALBLDFV	180	
DB	121	LIDQKITTSVRKLTALAGLQGLDGVYQKSLKNWLENNDTRARSVVVTVYIALBLDFV	180	

```
QY 181 AKIPSPAIQSGQVPLLSVVAQAANLHLLLRDASIFGAEWGFTPGBISTFYDQVTRTAQ 240
DB 181 AKIPSPAIQSGQVPLLSVVAQAANLHLLLRDASIFGAEWGFTPGBISTFYDQVTRTAQ 240
QY 241 YSDYCVKWTNTGLDKLKGNTAASWLKYHQFRREMTLLVLDLVALFPNYDTRTYPIETTTAQ 300
DB 241 YSDYCVKWTNTGLDKLKGNTAASWLKYHQFRREMTLLVLDLVALFPNYDTRTYPIETTTAQ 300
QY 301 LTRVYVTDPIVFNRETSGGFCRRWSLNSDISSEVESAVIRSPHLPDILSEIEFYTRAG 360
DB 301 LTRVYVTDPIVFNRETSGGFCRRWSLNSDISSEVESAVIRSPHLPDILSEIEFYTRAG 360
QY 361 LPLANNTEYLEYVGHSHIKYKNTNASSALERNYGTITSNKIKYDYLANKDIFQVRSIGADL 420
DB 361 LPLANNTEYLEYVGHSHIKYKNTNASSALERNYGTITSNKIKYDYLANKDIFQVRSIGADL 420
QY 421 ANYAQVYGVPIASFTLLDRNTGSGVGGFTYSKPHTTMQVCTQNTNTIDEIPPENEPLS 480
DB 421 ANYAQVYGVPIASFTLLDRNTGSGVGGFTYSKPHTTMQVCTQNTNTIDEIPPENEPLS 480
QY 481 RGVSHRLSHITSVSPSKNASSPARYGNLPPVATWHSADVNTVYSDKITQIPVVKARHL 540
DB 481 RGVSHRLSHITSVSPSKNASSPARYGNLPPVATWHSADVNTVYSDKITQIPVVKARHL 540
QY 541 VSGTTVIKPGPGFTGGNIIKRTSSGPLAYTSVSVKSPLSQRYRARIYASTTNLRLFWTIS 600
DB 541 VSGTTVIKPGPGFTGGNIIKRTSSGPLAYTSVSVKSPLSQRYRARIYASTTNLRLFWTIS 600
QY 601 GTRIYSINVNKTNKGGDILTFNTPDLATIGTAPFTSNYSDSLTVGADSPASGEVYVDF 660
DB 601 GTRIYSINVNKTNKGGDILTFNTPDLATIGTAPFTSNYSDSLTVGADSPASGEVYVDF 660
QY 661 ELIPVNATPEABEDLDVAKAVNGLFTSKDALQTSVTDYQVQNAANLVECLSDLYPNE 720
DB 661 ELIPVNATPEABEDLDVAKAVNGLFTSKDALQTSVTDYQVQNAANLVECLSDLYPNE 720
QY 721 KRLMDVAKEAKRLVQARNLLQDTGFNRINGENGTGSGTIEVAEGDVLFKDRSLRLTSA 780
DB 721 KRLMDVAKEAKRLVQARNLLQDTGFNRINGENGTGSGTIEVAEGDVLFKDRSLRLTSA 780
QY 781 REIDTETPTLYIQIDESLLKPYTRYKLGFIGSSQDLLEIKLIRHRANQIVKQVPDNLL 840
DB 781 REIDTETPTLYIQIDESLLKPYTRYKLGFIGSSQDLLEIKLIRHRANQIVKQVPDNLL 840
QY 841 PDVLPVNSCGGIDRCSEQOYVDANLALENNGENGNMSSDHAFTSHIDTGEIDLNENTGI 900
DB 841 PDVLPVNSCGGIDRCSEQOYVDANLALENNGENGNMSSDHAFTSHIDTGEIDLNENTGI 900
QY 901 WYVFKIPTTNGVATLGNLVBEGPLSGETLERAQQEQOQWQDKMARKEGASEKAYYAAK 960
DB 901 WYVFKIPTTNGVATLGNLVBEGPLSGETLERAQQEQOQWQDKMARKEGASEKAYYAAK 960
QY 961 QAIIDLFPADYQOKLNSGVENSDMLAAQNLVQSIPIVYNDALPEIPGMNTTFTLNTRL 1020
DB 961 QAIIDLFPADYQOKLNSGVENSDMLAAQNLVQSIPIVYNDALPEIPGMNTTFTLNTRL 1020
QY 1021 QQANWLYDLRNAIPNGDFNGLSDMNATSDVNQQLSDTSLVIPNWSQVSOQFTVQBN 1080
DB 1021 QQANWLYDLRNAIPNGDFNGLSDMNATSDVNQQLSDTSLVIPNWSQVSOQFTVQBN 1080
QY 1081 YRYVLRTARKGVGVDGYVIRIDGANQOTETLNFICDDDTGVLSDQTSYITKTVEFTPS 1140
DB 1081 YRYVLRTARKGVGVDGYVIRIDGANQOTETLNFICDDDTGVLSDQTSYITKTVEFTPS 1140
QY 1141 TQQWIDMSSETGVFNIESVELVLEE 1167
DB 1141 TQQWIDMSSETGVFNIESVELVLEE 1167
```

RESULT 2
US-10-032-717-4
; Sequence 4, Application US/10032717

```
; Publication No. US20020151709A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Feng
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn E. Sims
; TITLE OF INVENTION: Genes Encoding No. US20020151709A1el Proteins With
; Pesticidal Activity Against Coleopterans
; FILE REFERENCE: 35718/237005
; CURRENT APPLICATION NUMBER: US/10/032,717
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; TYPE: PR
; ORGANISM: Bacillus thuringiensis
; US-10-032-717-4
```

Query Match 54.0%; Score 3263.5; DB 13; Length 1210;
Best Local Similarity 53.4%; Pred. No. 3e-259;
Matches 657; Conservative 178; Mismatches 284; Indels 119; Gaps 21;

```
QY 1 MSPNQNEYEIILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSSEGENPELFGNP 60
DB 1 MSPNQNEYEIILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSSEGENPELFGNP 60
QY 61 ETPIS-SSVQVQIGIGVQVLAGALGVPPAGQIASFYFVQGLWPSSTVSVWEMIMKQVE 119
DB 61 ETVLSGQDAAKRAIDIVGKLSGLGVPPVPIVSLVTLQDILDLWPSQKQSEWEIFMEQVE 120
QY 120 DLIDQKITDSVKTALAGLQGLDGLDVYQKSLQWLENRNDTRASVVTQVIALDLP 179
DB 121 ELINQIKIAYARKALKALEGLGNNYQLYLTALKEWKNPNSGRALDRVNRREILDLSLF 180
QY 180 VAKIPSPAIQSGQVPLLSVVAQAANLHLLLRDASIFGAEWGFTPGBISTFYDQVTRTA 239
DB 181 TOYMSFRVTFNEFVPLVITVQANLHLLLRDASIFGAEWGFTPGBISTFYDQVTRTA 240
QY 240 QYSDYCVKWTNTGLDKLKGNTAASWLKYHQFRREMTLLVLDLVALFPNYDTRTYPIETTTAQ 299
DB 241 EYSDHCWKYETGLAKLKGTSKQVQVYQFRREMTLLVLDLVALFPNYDTRTYPIETTTAQ 300
QY 300 QLTREVYVTDPIVFNRETSGGFCRRWSLNSDISSEVESAVIRSPHLPDILSEIEFYTRTA 359
DB 301 QLTREVYVTDPIVFNRETSGGFCRRWSLNSDISSEVESAVIRSPHLPDILSEIEFYTRTA 356
QY 360 GLPLANNTEYLEYVGHSHIKYKNTNASSALERNYGTITSNKIKYDYLANKDIFQVRSIGLA 418
DB 361 S-ISSARVIRWAGHQISYHRIIPSDNIIKQVGTQNLHSTSTPDTNVDYIKTSLKDA 414
QY 419 DLAN-----YYAQVYGVPIASFTLLDQK--NTSGSGVGGFTYSKPHTTMQVCTQNTNTID- 470
DB 415 VLLDIVPFGTYTIFFGMEPEVEPFMVNQLANTRK-----TLKYNPVS 456
QY 471 -----RIPPE--NEPLSRGVSHRLSHITSVSPSKNASSPARYGNLPPVATWHS 517
DB 457 DIIAGTRDSELEPPETSDQPNYBSYHRLCHITSIPATGSGT-----GLVPFVSWTHRS 511
QY 518 ADVNTVYSDKITQIPVVKARHL---VSG---TTVIKPGFTGCGNLTSSGLAYTSV 571
DB 512 ADLINAHSKDTQIPVVKVSDLAISITGCPNNTVSGPGFTGGGIIKVRNG-VLIISM 570
QY 572 SVK-SPLSQRYRARIYASTTNLRLFWTISGTRIYSINVNKTNKGGDILTFNTPDLATIG 630
DB 571 RVKIGIDINKEYSMRIRYASANTFEYINPSEBENKVS-HAQKTNMRGEALTYNKFNATL- 628
QY 631 TAFTTSNYSDSLTVGADSPAS---GGEVYVDFKELIPVNATPEABEDLDVAKAVNGLPT 687
```



```

QY 1108 TETLTFNICDDDT-GVLSA----- 1125
DB 1103 TETLTFNSADYDTNGMYNTQVSTNGYNTNAYNTQASSTNGYNNMNTQASNTNGYN 1162
QY 1126 -----DQTSYITKTVEPTSPTEQVWIDMSETEGVNIESVELVLEE 1167
DB 1163 TNSVNDQGYITKTVEPTSPTEQVWIDMSETEGVNIESVELVLEE 1210

RESULT 4
US-10-606-320-4
; Sequence 4, Application US/10606320
; Publication No. US20040091505A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F. H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE OF INVENTION: Activity
; FILE REFERENCE: 35718/263948
; CURRENT FILING DATE: 2003-06-25
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1210
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-606-320-4

Query Match 54.0%; Score 3263.5; DB 16; Length 1210;
Best Local Similarity 53.4%; Pred. No. 3e-258;
Matches 667; Conservative 178; Mismatches 284; Indels 119; Gaps 21;

QY 1 MSPNNQNEYILDASSSTVSNSVRYPLANDQTTLLQNNYKDYLRMSEGENPELFCNP 60
DB 1 MSPNNQNEYIIDATPSTVSNSVNSRYPFANEPTNALQNMKYDKLKSAGNASEYPCSP 60
QY 61 EFTFIS-SSTVQIGIGVQVLTGALVPPFAGQIASPYSPVIGQLWPSSTVSVMIMKQVE 119
DB 61 EVLVSQDAKAAIDIVGKLLSGLGVPFVGPVIVSLYTQILDILWPSGQKQWEIPMEQVE 120
QY 120 DLIDOKITDSVRKTAGLOGLDGDLVYQKSLKNWLENDRTRARSVVVYQYIALELDF 179
DB 121 ELINOKIABYARNKALSELEGNNYQYLYLTALEWKENPNSRALRDVNRNRFELDLSLF 180
QY 180 VAKIPSPFISQVEPLLSVYAQAANLHLLLRDASIFGAEWGFTPGESTPTVDROVTRTA 239
DB 181 TQYMESFRVTPVEVPLVYVYQAAANLHLLLRDASIFGEWGWSTTINNYYDROKMLTA 240
QY 240 QYSDCYVKRYNTGLDKLGTNAASWLKQYQFREMTELLVLDLVALFPNYDTRTYPIETTA 299
DB 241 EYSDHCWKWYETGLAKGTSAKQWVQYQFREMTELLVLDLVALFPNYDTRTYMETKA 300
QY 300 QLTREVVYDPIVFNRETSGFCRWSLNSDTSFVSBSAVIRSPLDILSEIEPTTTRA 359
DB 301 QLTREVVYDPLGAVNVSSIG---SW-YDKAPSPGVIESVIRPPHPVYITGLTVYTSQR 356
QY 360 GLPLNNTYELEVWGHSTKYKNTNASSALERNYGTITS-NKIKYVDLANKDIFQVRSLSGA 418
DB 357 S--ISSARYIRHWAGHQISYHRIPSDNIKKOMYGNQNLHSTSTFDFTNYDIYKLTSLKDA 414

```

```

QY 419 DLAN-----YAAQVYGVYPVASFLLDK--NTGSGSVGGFTYSKPHHTMQVCTQNTYDID- 470
DB 415 VLLDIVFPQYTYIFGMEPEVSEFPMVQNLANTRK-----TLKNVPSVK 456
QY 471 -----RIPPE--NEPLSRGVSRLSHITSVSPSKNASSPARYGNLVPANWTHRS 517
DB 457 DIIAGTRUSELELPETSDQPNYESYSHLCHITSIPATGSTT-----GLVPVPSWTHRS 511
QY 518 ADVNTVYSDKITQIPVVKAHYTL-----VSG---TTVWKGPGPTGGNLLKRTSSGPLYTSV 571
DB 512 ADLINAHVSDKITQIPVVKVSLAPSTGGPNNTVVGSGPTGGGIIKVRNG-VIIISHM 570
QY 572 SVK-SPLSQRYRARIYASTTNLRLPVTISGRIRIYSINVKNTMKGDDTLTFNFDLATIG 630
DB 571 RYKISDINKYSMRIRYASANTTEFYINPSEENVKS-HAQKTMNREALTYPKFNATL- 628
QY 631 TAFTPSNYSDSLTVGADSPAS---GGEVVDKFEELIPVNATPEABEDLDVAKKANGLFT 687
DB 629 PPIKFTTTEPPTITGAIPEAEDFLGIEAYIDRIEIPVDETYEAEQDLEAAKANVALFT 688
QY 688 SKKDALQTSVTDYQVNOAANLVECLSDLYPNKRMMLMDAVKXAKLVOARNLQDTGPN 747
DB 689 NTKDGLRPGVTDYEVNQAAANLVECLSDLYPNKRELLFDVAREAKLSEARNLQDDPQ 748
QY 748 RINGENGWGTSGTGEVVAEGDVLFKDRSLRLTSAREIDTETTYTYLYQQIDSLKPYTRY 807
DB 749 EINGENGWGTASTGIEVIEGDLFKGRYLRLPGAREIDTETTYTYLYQYKEGVLKPYTRY 808
QY 808 KLKGFIGSSQDLEIKLIRHRANOIVKNVPDNLDPVLVNSCGGIDRCSEQQVVDANLAL 867
DB 809 RLRGFVGSQGLEIFTHHQNRIVKNVPDNLDPVLVNSCGSINRCSEQKYVNSRLEV 868
QY 868 ENNGENGMSSDSHAPSFIHDTEIDLANENTGIWVFKIPTTNGVATLGNLELVERGPLS 927
DB 869 ENR-----SGEAHEFSIPIDTEIDYENAGIWNVFKITDPEGVATLGNLELVERGPLS 922
QY 928 GETLERAQQQOQWQDQVARKGASEKAYAAKQAIIDLFPADYQOQKLSGVMSDMLAA 987
DB 923 GDALERLQREBQWQKIQMTRREBETDRRYMASKQAVRLYADYQOQLNPDEITDLTAA 982
QY 988 QNLVQSIPIVYNDALPEIPGMNNTSFTELTKRLOQAWNLVDLURNALPNDGPNGLSDMNA 1047
DB 983 QDLQSIPIVYVNEPPEIPGMNNTSFTELTKRLOQAWNLVDLURNALPNDGPNGLSNWNA 1042
QY 1048 TSDVNVQQLSDTSVLVIFPNNSQVSQFTVQPNRYVLRVLTARKEGVGDGYVIIRDGANO 1107
DB 1043 TPGVEVQIINTSVLVI PNMBEQVSQFTVQPNRYVLRVLTARKEGVGDGYVSIIRGQNO 1102
QY 1108 TETLTFNICDDDT-GVLSA----- 1125
DB 1103 TETLTFNSADYDTNGMYNTQVSTNGYNTNAYNTQASSTNGYNNMNTQASNTNGYN 1162
QY 1126 -----DQTSYITKTVEPTSPTEQVWIDMSETEGVNIESVELVLEE 1167
DB 1163 TNSVNDQGYITKTVEPTSPTEQVWIDMSETEGVNIESVELVLEE 1210

RESULT 5
US-10-032-717-2
; Sequence 2, Application US/10032717
; Publication No. US20020151709A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Feng
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn E. Sims
; TITLE OF INVENTION: Genes Encoding No. US20020151709A1e1 Proteins With
; FILE REFERENCE: 35718/237005
; CURRENT APPLICATION NUMBER: US/10/032,717
; CURRENT FILING DATE: 2001-10-23

```

```
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1206
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
; US-10-032-717-2

Query Match      53.7%; Score 3246.5; DB 13; Length 1206;
Best Local Similarity 53.4%; Pred No. 7.5e-257;
Matches 665; Conservative 172; Mismatches 290; Indels 119; Gaps 18;

Qy 1 MSPNNQNEYIILDASSSTVSNSVRYPLANDOTTTLQNNYKDYLRMSGENPELPGNP 60
Db 1 MSPNNQNEYIILDATPSTVSNSNRYPPANEPTNALQNMNDYKDYLRMSAGNASEYPGSP 60

Qy 61 EFTIS-SSTVOTGIGIVGVLGALGVPPAGQIASFSYFVQQLWPSSTVSWMIMKQVE 119
Db 61 EVLVSGQDAKAAIDIVGKLLGLGVFPVGPVIVSLYTLQILDLWPSGKSKQMBIFMEQVE 120

Qy 120 DLIDOKITDSVRKTALAGLQGLDGVYQKSLKNWLENRNDTRARSVVVTVQYIALELDF 179
Db 121 ELINQKIAETARNKALSELGLGNVYQLYLTALBEEENPNSRALRDVNRFEILDSLF 180

Qy 180 VAKIPSAISQGVFLSVYAAANLHLLLRDASIFGAEWGFTPGIBISTFYDQVTRTA 239
Db 181 TQMPSPFRVNTFVPEFLTVYMAANLHLLLRDASIFGEWGWSTTTINNYYDRQMKLTA 240

Qy 240 QYSDYCVKNTGDLKLGNTNAASWLKHYOPREMLLLDLVALPNYDTRTPYETTA 299
Db 241 EYSDHCWKVYETGLAKLGTSAKQWYDYNQFRREMLAVLDVVALPNYDTRTPYETRA 300

Qy 300 QLTREYVTDIVNRETSFGCFRRWSLNSDISPSEVESAVIRSPHFLDILSEIEFYTRA 359
Db 301 QLTREYVTDPLGANVSSIG--SW-YDKAPSPGVIESVIRPPHFDYITGLTVTQSR 356

Qy 360 GLPLNTEYLEYVWGHISIKYKNTNASSALERNYGTITS-NKIYDIANKDIFQVRSIGA 418
Db 357 S--ISSARIYRWAGHQISYHRVSRGSLQOMYGTQNLHSTSTFDFTNVDIYKTLSDA 414

Qy 419 DLAN-----YVAVYGVYASFTLLDK--NTGSGSVGGFTYKPHHTMQVCTQYNTID- 470
Db 415 VLADIVVPGYTYIFPGKPEVEFFWVNLNTRK-----TLKYNPVSK 456

Qy 471 -----ELPPE--NEPLSRGSHRLSHITSYSFSKNASSPARYGNLPFAWTHRS 517
Db 457 DIIASTROSELELPPEPSTQDNYESYSHRLCHITSIPATGNTI-----GLVPFVSWTHRS 511

Qy 518 ADVNTVYSKTIQIPVVAHTLVSGTIVKGPFGTCGNILKRTSS-----GPLAYTSVS 573
Db 512 ADLNTIYSKTIQIPAVKQWDLNLPFPVVKRGHGTGDLQYNRSTGSGVTLFLARYGL 571

Qy 574 KSPLSQRYRARIYASTNRLRFLVTISGTRIYSINVKNTWKNGDDLTFNTPDLATIGTAP 633
Db 572 ALEKAGKRYRVLRYATDADIVLVHND-----QIQPKTNWPGSDLTSTFKVADAITTL 626

Qy 634 ---TFSNVSDSLTVGADSPAS-CGEVVYVDKFLIPVNATFEAREDLQVAKVANGLPFSK 689
Db 627 NLATDTSALAKHNGEDPNSTLSGIVYVDRIEPIFVDDETYEREQDLAAKAVNALFTNT 686

Qy 749 KDALQTSVTVQVNOAANLVECLSDDELYPNKRLMDAVKAEKLVQARNLLQDTGFNRI 749
Db 687 KDGLRPGVTDYEVNOAANLVECLSDDLYPNEKRLLPDAVREKELSEARNLLQDPDQEI 746

Qy 750 NGENGWGTSIGIEVABGDVLPKDRSLRLTSAREIDTETPTLYLQQIDESLLKPYTRYKL 809
Db 747 NGENGWGTSIGIEVIEGDALPKGRYLRPLGAREIDTETPTLYLQKVEEGVLPKYTRYRL 806

Qy 810 KGFIGSSQDLIELKIRHANCIVKRVNPNLLPDVLVUNSCGIDRCSEQQVVDANTALEN 869
Db 807 RGFVSSQGLEIFTRIHTNRIIVKNVPDDLPLPDVSPVNSDGSINRCSEQKYVNSRLEVEN 866
```

```
870 NGENGNSSDSHAPSFHIDTGEIDLNENTGCIWVVKIPTTNGYATLGNLELVEBPLSGE 929
867 R-----SGEAHEFSIPIDTGEIDYENAGIWWGFKITDPEGYATLGNLELVEBPLSGD 920
930 TLERAQOQOQOQODKMARKGSEKAYAAKADRLFPADYQDQKLNAGVMSDMLAARQ 989
921 ALERLQREBQOQWKIQTREBETDRRYMASQAVDRLYADYQDQQLNPDPVITDLTAQD 980
990 LVQSIPIVYNDALPEIPGMNYTSFTLTNRLOQAANLYLDRNAIPNGDFRNLGSLWNATS 1049
981 LIQSIPIVYVNMPEPFGMNYTKFTLTDRLOQAANLYLDRNAIPNGDFRNLGSLWNATP 1040
1050 DVVVQQLSDTSVLVIPNNNSOVSOQFTVQBNRYVLTARKEGCGVGVYIIRDCANOTE 1109
1041 GVEVQQINHTSVLVIENWDEQVSOQFTVQBNRYVLTARKEGCGVGVYIIRDCANOTE 1100
1110 TLFNLCDDDT-GVLSA----- 1125
1101 TLFPSADYDTNGMYNTQVSTNGYNTNAYNTQASSTNGYNANMYNTQASNTNGYNTN 1160
1126 ----DQSYITTKTVETFEPTFEQWIDMSETEGVFNIESVELVLEER 1167
1161 SVYNDQGTITKTVPITPYPYTDQWMIENSETEGTFYIESVELIVDVE 1206

RESULT 6
US-10-414-637-2
; Sequence 2, Application US/10414637
; Publication No. US20030177528A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Feng
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn B. Sims
; TITLE OF INVENTION: Genes Encoding No. US20030177528A1el Proteins With
; TITLE OF INVENTION: Pesticidal Activity Against Coleopterans
; FILE REFERENCE: 35718/237005
; CURRENT APPLICATION NUMBER: US/10/414,637
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: US/10/032,717
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1206
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
; US-10-414-637-2

Query Match      53.7%; Score 3246.5; DB 14; Length 1206;
Best Local Similarity 53.4%; Pred No. 7.5e-257;
Matches 665; Conservative 172; Mismatches 290; Indels 119; Gaps 18;

Qy 1 MSPNNQNEYIILDASSSTVSNSVRYPLANDOTTTLQNNYKDYLRMSGENPELPGNP 60
Db 1 MSPNNQNEYIILDATPSTVSNSNRYPPANEPTNALQNMNDYKDYLRMSAGNASEYPGSP 60

Qy 61 EFTIS-SSTVOTGIGIVGVLGALGVPPAGQIASFSYFVQQLWPSSTVSWMIMKQVE 119
Db 61 EVLVSGQDAKAAIDIVGKLLGLGVFPVGPVIVSLYTLQILDLWPSGKSKQMBIFMEQVE 120

Qy 120 DLIDOKITDSVRKTALAGLQGLDGVYQKSLKNWLENRNDTRARSVVVTVQYIALELDF 179
Db 121 ELINQKIAETARNKALSELGLGNVYQLYLTALBEEENPNSRALRDVNRFEILDSLF 180

Qy 180 VAKIPSAISQGVFLSVYAAANLHLLLRDASIFGAEWGFTPGIBISTFYDQVTRTA 239
Db 181 TQMPSPFRVNTFVPEFLTVYMAANLHLLLRDASIFGEWGWSTTTINNYYDRQMKLTA 240

Qy 240 QYSDYCVKNTGDLKLGNTNAASWLKHYOPREMLLLDLVALPNYDTRTPYETTA 299
Db 241 EYSDHCWKVYETGLAKLGTSAKQWYDYNQFRREMLAVLDVVALPNYDTRTPYETRA 300

Qy 300 QLTREYVTDIVNRETSFGCFRRWSLNSDISPSEVESAVIRSPHFLDILSEIEFYTRA 359
Db 301 QLTREYVTDPLGANVSSIG--SW-YDKAPSPGVIESVIRPPHFDYITGLTVTQSR 356

Qy 360 GLPLNTEYLEYVWGHISIKYKNTNASSALERNYGTITS-NKIYDIANKDIFQVRSIGA 418
Db 357 S--ISSARIYRWAGHQISYHRVSRGSLQOMYGTQNLHSTSTFDFTNVDIYKTLSDA 414

Qy 419 DLAN-----YVAVYGVYASFTLLDK--NTGSGSVGGFTYKPHHTMQVCTQYNTID- 470
Db 415 VLADIVVPGYTYIFPGKPEVEFFWVNLNTRK-----TLKYNPVSK 456

Qy 471 -----ELPPE--NEPLSRGSHRLSHITSYSFSKNASSPARYGNLPFAWTHRS 517
Db 457 DIIASTROSELELPPEPSTQDNYESYSHRLCHITSIPATGNTI-----GLVPFVSWTHRS 511

Qy 518 ADVNTVYSKTIQIPVVAHTLVSGTIVKGPFGTCGNILKRTSS-----GPLAYTSVS 573
Db 512 ADLNTIYSKTIQIPAVKQWDLNLPFPVVKRGHGTGDLQYNRSTGSGVTLFLARYGL 571

Qy 574 KSPLSQRYRARIYASTNRLRFLVTISGTRIYSINVKNTWKNGDDLTFNTPDLATIGTAP 633
Db 572 ALEKAGKRYRVLRYATDADIVLVHND-----QIQPKTNWPGSDLTSTFKVADAITTL 626

Qy 634 ---TFSNVSDSLTVGADSPAS-CGEVVYVDKFLIPVNATFEAREDLQVAKVANGLPFSK 689
Db 627 NLATDTSALAKHNGEDPNSTLSGIVYVDRIEPIFVDDETYEREQDLAAKAVNALFTNT 686

Qy 749 KDALQTSVTVQVNOAANLVECLSDDELYPNKRLMDAVKAEKLVQARNLLQDTGFNRI 749
Db 687 KDGLRPGVTDYEVNOAANLVECLSDDLYPNEKRLLPDAVREKELSEARNLLQDPDQEI 746

Qy 750 NGENGWGTSIGIEVABGDVLPKDRSLRLTSAREIDTETPTLYLQQIDESLLKPYTRYKL 809
Db 747 NGENGWGTSIGIEVIEGDALPKGRYLRPLGAREIDTETPTLYLQKVEEGVLPKYTRYRL 806

Qy 810 KGFIGSSQDLIELKIRHANCIVKRVNPNLLPDVLVUNSCGIDRCSEQQVVDANTALEN 869
Db 807 RGFVSSQGLEIFTRIHTNRIIVKNVPDDLPLPDVSPVNSDGSINRCSEQKYVNSRLEVEN 866
```



```
Db 572 ALEKAGKTRVRLVATADADIVLHND-----QIMPKTINPGBDLTSKTPKVAADATTL 626
Qy 634 ---TFSNYSDSLTVGADSFAS-GGEVYVDKPELIPVNAATPEABEDLDVAKKAVNGLFTSK 689
Db 627 NLATDSSLAKHNLGEDPNSTLSGIVYVDRIEPIPVDETYEARQDLBAKAVNALPTNT 686
Qy 690 KDALQTSVTDYQVNOQANLVECLSDLELPNPKRMMLWDAKAKLVOARNLLQDTGENRI 749
Db 687 KOGLRPGVTDYEVNOQANLVECLSDLELPNPKRLLFDVAREAKLSEARNLLQDPQOEI 746
Qy 750 NGENGTGSGTGEVAGDVLFPKDSRLTSAREIDTETPTLYLQOIDESELLKPYTVKL 809
Db 747 NGENGTASTGIEVIEGDALEFGKRLPGAREIDTETPTLYLQKVEEGVLPKYTVRL 806
Qy 810 KGPTGSSQDLLEIKLIRHANOIVKXNPNLPLVLPVNSCGGIDRCSEQQOYVDANLLEN 869
Db 807 RGFVSSQGLEIPTRHQTNRIKVNQVDDLLPDPVSPVNSDCSINRCSEQKYVNSRLEVEN 866
Qy 870 NGENGNSSSHAFSPHIDGIEDLNENTGIWVFKIPTTNGYATLGNLELVERGPLSGB 929
Db 867 R-----SGEAHEFSPIDTGEIDYNNENAGIWWGFKITDPGEYATLGNLELVERGPLSGD 920
Qy 930 TLERAQOQEQOQDKMARKGASKEYAAKQADRLFPADYQDOKLNSGVEMSDMLAON 989
Db 921 ALERLQEEQOMKQMTTRREEDTDRYNASKQAVDRLYADYQOQQLPDPVEITDLTAQD 980
Qy 990 LVOSIPVYNDALPEIPGMNYTSTELTNRLQQAANLYDLRNLAPNGDFRNLGSDMNATS 1049
Db 981 LIQSIPIVYNEPPIPGMYTKTELTDRLQQAWSLYDQBNALPNGDFRNLGSDMNATP 1040
Qy 1050 DVNVQQLSDTSVLVIPNNNSQVSOQFTVQPNRYRVLRTARKEGVDGYVIIRGANOTE 1109
Db 1041 GVEVQIINHSTSVLIPNWDQVSOQFTVQPNRYRVLRTARKEGVDGYVSIIRGQNOTE 1100
Qy 1110 TLTFNICDDDT-GVLSA----- 1125
Db 1101 TLTFASDYDTNGYNTQVNTNGYNTNAYNTQASSTNGYNNANMYNTQASNTNGYNTN 1160
Qy 1126 ---DQSYITKIVETPSTEQVIMDSMETEGVFNIESVELVLEE 1167
Db 1161 SVYNDQGYITKIVTFPIYTDQWIMSETEGTYIESVELIVDVE 1206
```

RESULT 8

US-10-099-285-72
: Sequence 72, Application US/10099285
: Publication No. US20030105319A1

GENERAL INFORMATION:

APPLICANT: Schnepf, H. Ernest
: Wicker, Carol
: Narva, Kenneth E.
: Walz, Michelle
: Stockhoff, Brian
: Muller-Cohn, Judy
TITLE OF INVENTION: Toxins Active Against Pests
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSER: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/099,285
FILING DATE: 15-Mar-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/002,285
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/886,615
FILING DATE: 1-JUL-1997
APPLICATION NUMBER: US 08/674,002
FILING DATE: 1-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-701C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 1156 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-10-099-285-72

Query Match 51.3%; Score 3100.5; DB 14; Length 1156;

Best Local Similarity 52.0%; Pred. No. 6.9e-245;

Matches 615; Conservative 196; Mismatches 330; Indels 41; Gaps 13;

```
Qy 1 MSPNNQNEYELTDASSSTVSVDNSVRYPLANDQTTTLQNNMYKDYLRMSEGNPELFGNP 60
Db 1 MNRNNQNEYELIDPHGCCPSDDVRIPPLASDPNAALQNNMYKDYLMQTDDEYTSYNP 60
Qy 61 ETFTISS-STVQTGIGIVGVLGALGVPPAGQIASFYSFVQGLWPSSTVSVMEMKQVE 119
Db 61 SLSISGRDAVQALTUVGRILGALGVPPSGQIVSFYQFLATLWPNVDTAIWEAPMQVE 120
Qy 120 DLIDQKITSVRKTAALAGLQGLGVYQKSLQNLWLRNDTRARSVVATQYIILELDP 179
Db 121 ELVNNQITFARNQALRQLGLGDSFNYYQSRSLQNLWLRNDTRNLVSVRAQFIALLDLP 180
Qy 180 VAKIPSPAIQGOEVPILSVVAAQANLHLLLRDASIFGAEWGFTPGHISTFYDQVTRTA 239
Db 181 VNAIPFAVNGQVPIILSVVAAQVNLHLLKDSLFGEGGFTQGEISTFYDQLELTA 240
Qy 240 QYSDYCVKWNVTGLDKLGTNNAASWLKYHOPRREMTLLVLDPNYPNRYPIETTA 299
Db 241 KYTNYCETWNTGLDLRGNTESWLRYHQFRREMTLLVLDVVALFFYDVELYPTGSPN 300
Qy 300 QLTREVYTDPIVFNRETSGGFCRWLSNDSISPSEVSATVRSRPHLPDLSEIBFYTTTA 359
Db 301 QLTREVYTDPIVFNPPANVGLCREWGTNPYNTFSELENAPIRPHLPDLRLSLTSSNR- 359
Qy 360 GLPLNNTLEYLWVGHSHIKYKNTNASSALERNYGTITSNKIKY---YDLANKDIFQVRS 416
Db 360 -FPV-SSNPFMDYWSGHTLRRLSYLNDSAVQEDSYGLITTYTRATINPGVDGTHR----IEST 413
Qy 417 GADLANYYAIVGYVPYASFTL--LDKNVTGSGVSGFTYKPHHTMQVCTQNYNTIDEIPP 474
Db 414 AVDPFRSALIGIYGVNRASFVPGGLPNGTTSFANGG-----CRDLYDTNDELPP 461
Qy 475 ENEPLSRGYS-HRLSHITSYSPSN-ASSPARYGNLVPVANTHRSDVNTVYSDKIQT 532
Db 462 DB---STGSSTHRLSHVTPFSTQNOAGSIANAGSVPTVYVWTRRDVDLNNITTPRITQL 518
Qy 533 PVKKAHTLVSGTIVIKGPGFTGMKLTKRTSSGLAYTTSVKSPUSQRYARIRVASTTN 592
Db 519 PLVKASAPVSGTIVLKPGPGTGGGLRNTNGTGTGLVTVNSPLTQYRVRVRPSSGN 578
Qy 593 LRLPVTISGTRIYSINVNKNKMGDDLTFTNFDLATICT-----APTFNSYSDSLTVGAD 647
Db 579 PSIRILRGNTSIAVQRFSGSTNRCGELTYESFVTSEPTTNGSDLPFTTQAOENLTILAE 638
Qy 648 SPASGGVYVDKPELIPVNATFEABEDLDVAKKAVNGLFTSKDQALQTSVTDYQVNOQAN 707
```


RESULT 10
US-10-614-524-6
; Sequence 6, Application US/10614524
; Publication No. US20040016020A1
; GENERAL INFORMATION:
; APPLICANT: Arnaut, Greta
; APPLICANT: Boets, Anemie
; APPLICANT: Damme, Nicole
; APPLICANT: Mathieu, Eva
; APPLICANT: Vanneste, Stijn
; APPLICANT: Van rie, Jeroen
; TITLE OF INVENTION: Insecticidal proteins from *Bacillus thuringiensis*.
; FILE REFERENCE: NEWBTSUS2
; CURRENT APPLICATION NUMBER: US/10/614,524
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: US/09/739,243
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/173387
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: *Bacillus thuringiensis*
US-10-614-524-6

Query Match 47.78; Score 2884.5; DB 15; Length 1152;
Best Local Similarity 49.14; Pred. No. 3.8e-227;
Matches 578; Conservative 200; Mismatches 365; Indels 33; Gaps 14;

Qy 1 MSPNNQVEYELDASSSTVSNSVRYPLANDQTTTLQNNNYKDYLRMSGCEPNELFGNP 60
Db 1 MNRNNDYEVIDASNGCCASDDVQYPLARDPNAVFQNMHYKDYLTQYDGYDYGSPNP 60

Qy 61 ETPISSSTV-QTGIGIVQVGLGALGVPFAGQIASFYFVGLWPSSTVSWEIMKQVE 119
Db 61 NLSINPRDVLQGTGINIVGRLLGFLGVPFAGQVLTFTYLLNQLWPTNDNAWAEAFMAQIE 120

Qy 120 DLIDOKITDSVRKTAGLGLDGLDVYOKSLNMLENENDTFAARSVVVTOYIALRDP 179
Db 121 ELINQRISEAVGTAADHLTGLHDNYELVEALESEWLERPNAKT-NLLFNRTFLDLSLF 179

Qy 180 VAKIPSPA1----SGQVPLLSVYAQAANLHLLLRDASIFGAEGFTPGESTFYDRQV 235
Db 180 TQMPSPGTGPGSQNYAVPLTVYAQAANLHLLLRDASIFGAEGFTPGESTFYDRQV 239

Qy 236 TRTAQYSDYCVKQNTGLDKLGTNAASWLKYHQFREMILLVLDLVALPNNYDRTYPI 295
Db 240 ERTQYTNHCVTYNTGLDLRGNTESWLNHYHFRREMTLWMDLVALPNNYDRTYPI 299

Qy 296 ETQAOLTRVYTDIVFNRETSQFCRWSLNSDISSEVESAVIRSPHLDLILSEIFY 355
Db 300 GANPOLTREITDPIVYNPPANQICRRWGNPNYTFSELENAPIRPHLPDLRLNRTIS 359

Qy 356 TTRAGLPLNNTYLEYVWGHISIKYKNTNASSALERNYGTITSNKIKYVDLANKDIFOVRS 415
Db 360 RNRYPATPNS-YLDYWSGHLLQOYANNPTYTSYQITSN-TRLFNTN-CANAIDS 416

Qy 416 LGADLANVYQVGPVPSFTLLDKNTGSGVGGFTYSKPHTTMOCVTQNYNTIDEIPPE 475
Db 417 RARNFGLNLYANLYGVSYLNI-----PFTGVMSBIT-SAPWT----CWQDLTTEELPLV 465

Qy 476 NEPLSRGSHLSHITSYSPSKNASSP-ARYGNLPFWAHTRSADVTNVTYSDKITQIPV 534
Db 466 NNPN-----NLLSVTLFRLNTTQGGPLATVGPFTVTVTRQDVFNNIITPNRTQIPV 520

Qy 535 VKAHTLVSGTIVKPGFTGNNILKRTSSGPLATVSVKSPLSQRYRARIYASTNLR 594
Db 521 VKAYELSSGATVWVGPGTGGDVIRRTNTGFGAIRSVTGLQRTAIRPRIYASTIDPD 580

Qy 595 LFTVTSITRIYSINVAKTNKGGDLTFTNTFDLTATIGTFTSNTSYSDSLTUCADSPASGGE 654
Db 581 FFVTRGTTINNFRFTNMRGQESRYESYRTVEFTTFFNFQSQDIIIRTSIQLSGNGE 640

Qy 655 VYVDKFEELIPVNATFRAEEDLDVAKAVNGLFTSKDALQTSVTDYQVQANLVCECLSD 714
Db 641 VYLDRIEIIIPNPTREAEEDLEAKKAVASIFTRDGLQVNVTDYQVQANLVCECLSD 700

Qy 715 ELYPNERGMLWDVAKBAKRLVQARNLLQDQFNRING--ENGWGTSGTGIEVAEGVLPKD 772
Db 701 EQYAHDKKMLLEAVRAAKRLSRERLLQDPDFTNTINSTENGWTKASNGVTISEGGPFYKG 760

Qy 773 RSLRLTSAREIDTETPTLYLQQIDESLLKDYTRYKLGKPGSSQDLLEIKLIRHRANQIV 832
Db 761 RALQLASAR----ENYPTIYQKVDASELKPTRYRLDGFVKSSQDLLEIDLIHHRKHLV 816

Qy 833 KNPVDNLLPDVLPVNSCGGIDRCSEQQYVDANLALENNGB--NGNNSDSHAFSPHIDTGE 891
Db 817 KNVLDNLVSDTYDDSCSGINRCCEQQMVAQLETEHHHPMDCCCEAAQTHEFPSSVIDTGD 876

Qy 892 IDLANENTGIWVPKIPPTNGYATLGNLLEVBEGPLSGETTLERAAQQOQOQWQDKMARKEGA 951
Db 877 LNSTVDQGIWVPKIPPTNGYATLGNLLEVBEGPLSGETTLERAAQQOQOQWQDKMARKEGA 936

Qy 952 SEKAYYAAQAIIDRLPADYQOKLNSGVHSDMLAAQNLVQSIPIVYNDALPEIRGMNYT 1011
Db 937 TDRVYQDAKQSIINHLFVDYQDQQLNPQIGHADIMDAQNLVASISDVVYSDAVLQIPGINYE 996

Qy 1012 SPTELTNRLQOAVNLVLRNAIPNGDFRNLSDMNATSDVNVQQLSDTSVLVIPNWSOV 1071
Db 997 IYIELSNRLQOAVNLVLRNAIPNGDFRNLSDMNATSDVNVQQLSDTSVLVIPNWSOV 1056

Qy 1072 SQQFTVQPNRYVLRVTRARKEGVDGYVIRIDGANQOTETLTFTNTCDDDTGVLASADQTSYI 1131
Db 1057 SQQFRVQPNCKYVLRVTRARKEGVDGYVIRIDGANQOTETLTFTNTCDDDTGVLASADQTSYI 1116

Qy 1132 TKTVETPSTEQWIDMSSETGVNIESVELVLEE 1167
Db 1117 TKCVIFYSHTERHMYEVNTEGAFHIDSIEPVEYETER 1152

RESULT 11
US-10-099-285-74
; Sequence 74, Application US/10099285
; Publication No. US20030105319A1
; GENERAL INFORMATION:
; APPLICANT: Schnepf, H. Ernest
; Narva, Kenneth E.
; Walz, Michelle
; Stockhoff, Brian
; Muller-Cohn, Judy
; TITLE OF INVENTION: Toxins Active Against Peste
; NUMBER OF SEQUENCES: 105
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/099,285
; FILING DATE: 15-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/002,285
; FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/886,615
FILING DATE: 1-JUL-1997
APPLICATION NUMBER: US 08/674,002
FILING DATE: 1-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-701C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 74:
SEQUENCE CHARACTERISTICS:
LENGTH: 1150 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 74:
US-10-099-285-74

Query Match 46.7%; Score 2825.5; DB 14; Length 1150;
Best Local Similarity 48.6%; Pred. No. 2,6e-222;
Matches 574; Conservative 197; Mismatches 364; Indels 47; Gaps 17;

QY 1 MSPNNQNEYILDASSSTSVSDNSVRYPLANDQTTLLQNMNYKDYLRMSGE-----NP 54
DB 1 MNRNNPEVEIIDAFCGCPDSDDDVRYPLASDPNAAAFQNMNKBLYQYDGDYTGSLNP 60

QY 55 ELFGPPEFISSTVGTGIGIVGVGLGALGUPPAGQIASFYSFVGLQWPSSTSVSWEMI 114
DB 61 NLSINPR-----DVLQGTINIVGRILGFLGVPFAGQLVTFYTFLLNQLMPTNDNAWEAF 115

QY 115 MKQVEDLIDOKITDSVRKTAGLGLGDLGVQKSLKWLLENRNDTRARSVVVTVIA 174
DB 116 MAQIEELIDOKISAGVVRNLDLGLHDYEEYLALEWLERENGARA-NLVTQREN 174

QY 175 LELOPVAKIPSPAIS-GQE---VPLLSVYAQAANLHLLLRDASIPGAEGFTFGEISTF 230
DB 175 LHAPVTRMPSFGTGPQSDQDAVALLVYAQAANLHLLLRDARIYGARWGLQOQINLY 234

QY 231 YDRQVTRTAQSDYCKVYNNGLDKGNAASMLKXHOFRMTLLVLDLVLPVYDT 290
DB 235 FNAQQRTRITYNHCVETINRGLDVRGTESMLNHRFRMTLLMAMDLVLPVYNNV 294

QY 291 RTYPLETTAQLTRKYVTPIVFNRETSGFCRRMSLNSDISFYSVESAVIRSPHLDLIS 350
DB 295 ROYPNGANPQLTRREYTPDIVNPPANQICERWGNPNYTFSELENAPIRPHLPERLN 354

QY 351 EIEFTYTRAGLPLNNTEYLEYVWGHISIKYKTNASSALERNYGTITSNKIKYDLANKDI 410
DB 355 RLTISSRNYTAPTNS-FLDYWSGHTLQSQHANNPTVETSYGQITSN-TRLENTTN-GA 411

QY 411 FQVRSGLADLANVYAQVGVYPASFTLLDXTGSSGVGFTYYSKPHITMQVCTQNYTID 470
DB 412 RAIDSRAARNFGLYANLYGV--SSLNIPP--TG-----VMSEITNAANTCRODLTTE 460

QY 471 BIPPNEPFLSRGSHRLSHITSYSFKNASSP-ARYGNLPPFANWTHRSADVNTVYSDKI 529
DB 461 ELPLENNP-----NLLSHVFLRNTYQGGPLATLGFVPVYVWTRVEDVFNITADRI 515

QY 530 TQIPVVKATLVSGTIVKGRFGTGNILKRTSSGGLAYTSVSKSPLSQRYRARIYAS 589
DB 516 TQLPVWKASEIGGTVVKGFGTGGDILRRTDGGAVGTIRANVAPLTPQYRILRYAS 575

QY 590 TTN--LRFLPVISGRIRIYSINVNKMGDDLTFTFPDLATIGTAPTNSVSDSLTVGAD 647
DB 576 TTSFVNLFVNNSAA---GFTLPSTMAQNGSLTYESFNTLEVTHYIRFSQSDTLRLNIF 632

QY 648 SPASGEVYVDKFLIPVNATPEAEEDLDVAKAVNGLPTSKDALQTSVTDYOVNOAAN 707
DB 633 PSISQGEVYVDKLEIVIPNTPREAEEDLEDKAVASIFTRDGLQVNTDYQVDOAAN 692

QY 708 LVBCLSDBLYPNKEMLDWDAVKEAKRLVQARNLLQDTGFNRING--ENGWTCSTGIEVAL 765
DB 693 LVSCLSDBQYGHDKMLLEAVRAAKRLSRERLLQDDPDFNEINTEENGWAKASNGVTISE 752

QY 766 GDVLPKDRSLRLTSAREIDTETYPVLYQOIDESESLKPYTRKLGKFGFSGSDLEIKLIR 825
DB 753 GGPFFKGRALQASAR----ENYPYIYQKVDASTLKPYTRKLGKFGVQSSQDLEIDLH 808

QY 826 FRANQIVRNVPDNLPLPVLVNSCGIDRCSEQQVYDANLALENNNGENGNMSSDSHAFSP 885
DB 809 EHKVHLVKNVPDNLVSDTYSDGSCSGINRCBQHQVDVQLDAEDHPKDCCEAAQTHEPSS 868

QY 886 HIDTGEIDLNTGTVVVKIPPTNGYATLGNLELVBESGPLSCETLRAQQOQEQQWQDM 945
DB 869 YIHTGDLNASVDQGIWVVLQVTRTDDGYATLGNLELVBEGPLSGESLEBQRDNKWNBSV 928

QY 946 ARKEGASEKAYAAKQAIIDLRFADYQDQKLSGVMSDMLAAQNLVQSPYVYNDALPEI 1005
DB 929 GRKAETDRIYQDAKQAINHLFVDYQDQQLSPFEGMADIIDAQNLIASISDVYSDAVLQI 988

QY 1006 FGMNYSFTSLTNRLQQAANLYDLRNALPBGDFRNLGLSDMNATSDVNVQQLSDTSVLVIP 1065
DB 989 FGINVEMTSLNRLQQAASLYLTSRNVVQNGDFNSGLDSMNATTDVAVQDGNMHPVLIS 1048

QY 1066 MNSQVSOQFVQPNRYVRLVARTAKSGVGGYVIRDCANQETETLTNLCDDDTGVLSA 1125
DB 1049 HMDAQVSOQFVQPNCKYVLRVARTAKKVGNGDGYVITQDGAHHRBTLTTFNACDYDVNGTHV 1108

QY 1126 DOTSVIITKVTFTSTROVWIDMSSETGVFNIESVELVEE 1167
DB 1109 NDNSTIITKELVFPYKTEHMMWVEVSETEGTFYIDSIEPIETQE 1150

RESULT 12
US-10-099-285-76
; Sequence 76, Application US/10099285
; Publication No. US20030105319A1
; GENERAL INFORMATION:
; APPLICANT: Schnepf, H. Ernest
; Wicker, Carol
; Narva, Kenneth E.
; Walz, Michelle
; Stockhoff, Brian
; Muller-Cohn, Judy
; TITLE OF INVENTION: Toxins Active Against Pests
; NUMBER OF SEQUENCES: 105
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/099,285
; FILING DATE: 15-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/002,285
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/886,615
; FILING DATE: 1-JUL-1997
; APPLICATION NUMBER: US 08/674,002
; FILING DATE: 1-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: MA-701C2

TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 1134 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 76:
US-10-099-285-76

Query Match 44.8%; Score 2710; DB 14; Length 1134;
Best Local Similarity 48.1%; Pred. No. 7.9e-213;
Matches 559; Conservative 195; Mismatches 345; Indels 62; Gaps 19;
Qy 39 NMNYKDYLRMSEGENPEL-----FGNPETFISSSTVGTGIGVQVILGALGVPAGQIA 92
Db 4 NPNINECIYPNCLSPNEVEVLGGERGNVT-----GLQTGIDIVAVVVGALGPPVGGILIT 58
Qy 93 SFYSPIVQQLWPSSTVSVMIMKQVBDLIDQKITDSVRKKTALAGLGGLDGLDVIYKSL 152
Db 59 GFLSTLFCGLWPSNDQAWAEAFIEQMBELIEQRISDQVVTALDDLTGIGNYNYQYLIAL 118
Qy 153 KRWLENRNDTRARSVVVTVQYIALELDFVAKIPSPAISGO-----EVPILLSVYQAANLHL 207
Db 119 KEWEERPNGVRA-NLVQRFELHALFVSSMPFSG-SGPGSORFOQLLVAVYQAANLHL 176
Qy 208 LLLDASIFGAEWGFTPEISTFTYDROV-TTAQYSDYCVKWTNTGLDKLGTWASWLK 266
Db 177 LLLADAKEYGARWLRESOIGNLYFNELOTRDYTNHCVNAYNGLAGLRTSABSWLK 236
Qy 267 YHQFERETLLVDLVALFPNVDTRYPIETTAQLTREVTYDPIVFNRETSGP----CR 322
Db 237 YHQFERETLLAMDLIALFPYNTERYPIAVNPQLTREVTYDPLGVPSSESLPELACL 296
Qy 323 RWSLNSDISFEVSASVIRSHPLFDLSEIFBYTTRAGLPLNNTLEYLWVGHSI--KYK 380
Db 297 RWOQTSAMTFNLENALISSPHLFDTLNNLMYTGSPFVHLTN-QLIBCNIGHSVTSLL 355
Qy 381 NTNASSALERNYGITNSKIKYDLDANKDIEQVRS---LGADLANVYQVYGVVPASFTL 437
Db 356 ASGPTTLVRRNYGSTTS-IVNYSFNDRDVYQINRSHTGIGFQN--APLFGITRAQF-- 410
Qy 438 LDKNTGSGVGFTYSKPHTTNQVCTQNYNTIDEIP--PENEPISRGYSHRLSHITSYSF 495
Db 411 -----YFGIYSVTQRNALTCEQYNSIDELPSLDNPEPISRSYSHRLSHITSYLH 461
Qy 496 -----SQNASPARYGNLVPFAMTHRSADVNTVYSDKITQIPVVKHATLVSGTIVYK 549
Db 462 RVLTIDGINIYS-----GNLPTVWVTHRDVLDLNTITADRTQLPLVKSFBIPAGTIVVRG 517
Qy 550 PGFTGNNILKTSGLSSPLATYSVKSPISORVYRIRVASTNLRFLTISGTRIYSINV 609
Db 518 PGFTGGDILRRTGVGTFTIIRVTAPUTQYRIRFPASTNLPFGIRVGDQRQVYDFP 577
Qy 610 NKTWNKGGDLTFTFDLATIGTAPTFNSYSLTVGADSFASGGEVYVDKPELIPVNAATP 669
Db 578 GRVWNRGDELAYESFATREFTTDFNFRQPELISVPANAFSAGQEVYFDRIEILIPVNPAP 637
Qy 670 EABEDLVAKVAVNGLFTSKKDALQTSVTDYQVQNAANLVCLSDLEYPNFKRMLMDAVK 729
Db 638 EAKEDLEAAKAVASLFTRTDGLQVNVKDYQVQNAANLVCLSDLEQYGDYKQMLLEAVR 697
Qy 730 EAKELVQARNLLQDTGENRING--ENWGTCSTGLEVAEGDVLFKDRSLRLTSAREIDTET 787
Db 698 AAKLSRERNLLQDPDFNTINSTENGWKAANGVTISEGSPFYKGRALQLASAR-----EN 753
Qy 788 YPTLYQOIDSLKPYTRYKLGKPGISSQDLKLRHRANQIVKNPVNDLLPDVLPVN 847
Db 754 YPTTIYQKVDASLKPTRYRSKDFVKSQDLLEIDLHKKVHLVKNVNDLVSDTTPDD 813

Qy 848 SCGGIDRCSEQQYVDANLALENGEB-NGMSSSDSHAPSHIDTGEIDLNTGNTGIWVVPKI 906
Db 814 SCGGINRCQEQQVNAQLETEHHHPMDCCBAAGTTFESSYIDTGLDNLSSVDQGIWAPKV 873
Qy 907 PPTNGYATLIGNLLEVEEPLSGETILERAQQOQOQOQDKMARKGASEKAYAAKQAIIDLRL 966
Db 874 RTTDTGATLIGNLLEVEEPLSGESIEREQRDNTKWSAELGRKRAETDRVYQDAKQGINHL 933
Qy 967 FADYQOQKLSGVEMSDMLAAQNLVQSIPYVYNDALPEIPGMNYTSFTLTLNRLQQAANL 1026
Db 934 FVDYQOQNLNPEIGMADIMDAQNLVASISDVYSDAVLQIPGINYEIYTTLSNRLQQAAYL 993
Qy 1027 YDLRNLAINPGRNGLSDWNATSDVNVQQLSTSVLVIWNASQVSQQTFTVQNTYVYLR 1086
Db 994 YTSRNVQNGDFNGLSDWNATAGASVQDQGNTHFLVLSHMDAQVSQQFRVPQNCYVLR 1053
Qy 1087 VTARKEGVGDGVVIIRDGANOTETITFNICDDDTGVLSADQTSYTKTVEFTPTSTEQVWI 1146
Db 1054 VTAERKVGGDGVVIIRDGAHTEITLTFACDYDINGTYVTDNTYLTKEVIFYSHTEHWW 1113
Qy 1147 DMSETGVFNIESVELVLEEE 1167
Db 1114 EWNTEGAFHIDSIEFVETEK 1134
RESULT 13
US-09-988-462-7
; Sequence 7, Application US/09988462
; Publication No. US20030046726A1
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; Desai, Malini M.
; Lewis, Kelly S.
; Kramer, Vance C.
; Warren, Gregory W.
; Byola, Stephen V.
; Crossland, Lyle D.
; Wright, Martha S.
; Merlin, Ellis J.
; Launis, Karen L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Syngenta Biotechnology, Inc.
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/988,462
; FILING DATE: 20-NO. US20030046726A1-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/547,422
; FILING DATE: 11-APR-2000
; APPLICATION NUMBER: US 08/459,504
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: S-188051
; TELECOMMUNICATION INFORMATION:


```
Qy 241 YSDYCVKWTGTLKGLKGTNAASWLVKQHPREMTLLVLDLVALFPNPDTRTYPIETTAQ 300
Db 227 YSDYCARWYNTGLANNLGRCTNAESWLVYQFPREDLTGLVLDLVALFPNPDTRTYPIETTAQ 286
Qy 301 LTRVYVTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSHPLFDILSBIEFYTTTAAQ 360
Db 287 LTRVYVTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSHPLFDILSBIEFYTTTAAQ 345
Qy 361 LPLANTVEYVWGHSHIKYKNTNASSALERNYGTITSNKIKYDVLANKDIFQVRSGLGADL 420
Db 346 -RWSNTQYMWYVGHSHIRSRIRGSLSTWTHGNTNTSINPVLTQTSRDVYRTSPAGIN 404
Qy 421 ANYTAQVGVYVYASFTLLDKNGTSGSGVGGFTYKSPHTMQVCTQNYNTIDIPPE--NEP 478
Db 405 ILLTPVNGVPWAFNWM--RNPLNSLRGSLLYTGYT--GVGTQLFDSLETPEPTTERP 460
Qy 479 LSRGYSHRLSHITSYSPSKNASSPARVGN---LVPFANTHRSADVNTVYSDKTIQIPV 535
Db 461 NYESYSHRLSNIRLIS-----GNTLRAPVYSWTHRSADRTNTISSDITQIPLV 509
Qy 536 KAHTLVSGTIVIKGPGFTGGMILKRTSSGPLAYTSVSVKSPLOKRYARIRYASTNMLRL 595
Db 510 KSFNLSGTSVSGPGFTGGDIIRTNVNGSVLSGMLNPNNTSLQRYRVRVRIASQTMVL 569
Qy 596 FVTISGTRIYSINVKNTWKGDLLTFMTFDLATIGTAFPSNYSDSLTVGADSPASGGEV 655
Db 570 RVTYVGGSTPDQGPSTMSANESLTSQSFRFAEPVPGISASG-SOTAGISISNAGRTQ 628
Qy 714 YVOKFELIPVNAATEAREDLVAKVANGLTSSKDD-ALQTSVTDYQVQNAANLVECLSD 714
Db 629 HFDKIEPIPIATPEABYDLERAQEAVALNTNTNPRLLKTGVTDYHIDEVSNLVACLSD 688
Qy 754 ELYPNEKRLMDAKVKAELVQANLLQDTCFNRI-----GENG 754
Db 689 EPLCDREKRELBKVKYAKRLSDERNLLQDPNFTSINKQPDPSNNBSGNFTSIHQSEHG 748
Qy 814 WTGSGTGEVAGDVLFDKRSRLTSAREIDTETPTLYQOIDSLLKPYTRYKLGPIG 814
Db 749 WNGSENITIQBGNDVFENYVTLPGT---FNECPTLYQKIGABLKAVTRYQLSGVIB 805
Qy 870 SSQBLEIKLIRHRAQIVKQNP--DNLLPQVL--PVNSCGGIDRCSEQQVVDANALENN 870
Db 806 DSQBLEIYLIRYNAKHETLDVPGTGESVWPLSVESPIGRGCPNRCA--PHEWNPDLDCS 863
Qy 929 GENG-NWSSDSHAPSFIHDTGEIDLNENTGTWVVPKIPPTNGYATLGNLSEVERGSLGB 929
Db 864 CRDGEKCAHSHHSLDIDVGCIDLHENLGVVWVFKIKTQGHARLGNLEFISEKPLLGE 923
Qy 989 TLERAQOQOQOQOQOQMARKGASBKAYAAKQAIIDRLFADYQDQKLSGVMSDMLAQN 989
Db 924 ALSRVKRAEKKWKREKQLKLETRKRVYTEAKEAVDALFVDSQYDRQLQADTWIGMHAADK 983
Qy 1049 LVQSIPTVYNDALPEIPGANYTSTELTNRLQQAANLYDLNAILPNGDFRGLSDWATS 1049
Db 984 LVHRIEAYLSLSVIFGVNABIFEELEGRITITISLYDAENVVYKNGDFNNGLACWVK 1043
Qy 1109 DVNVQQLSDTSVLVYPNNSVSOQFTQVQNYRVLRVYARKEGVGDGYVLIIRDCANOTE 1109
Db 1044 HVDVQSHHRSVLVPEWAEVSQVRVCPGCGYILRVYATKEGVGGCVTIHIEENND 1103
Qy 1129 TLTFNICDD-----DTGLVS-----ADQTS----- 1129
Db 1104 ELKPKNCEBEBVPTDGTGTCNDYTAHQCTAVCNSERNAGYEDAVEVDVTTASVNYKPTVEE 1163
Qy 1130 -----YTKVTETPTEBQWIDMSHTEGVNFIESVELV 1163
Db 1164 TYTDVRDNHCEYDGYVYPPVPAGYMTKELEYFPETDKVWIEBIGETEGKFIIVDSVELL 1223
Qy 1164 LEE 1167
Db 1224 LME 1227
```

```
RESULT 15
US-09-826-660-23
; Sequence 23, Application US/09826660
; Patent No. US2001002690A1
; GENERAL INFORMATION:
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Stelmean, Steven J.
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
; FILE REFERENCE: MA-714XC2D1
; CURRENT APPLICATION NUMBER: US/09/826,660
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/178,252
; PRIOR FILING DATE: 1998-10-23
; PRIOR APPLICATION NUMBER: 60/065,215
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/076,445
; PRIOR FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 1186
; TYPE: EXT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Toxin encoded by synthetic B.t. gene
US-09-826-660-23
```

Query Match 40.6%; Score 2454.5; DB 9; Length 1186;
Best Local Similarity 43.1%; Pred. No. 8.3e-192;
Matches 536; Conservative 191; Mismatches 382; Indels 135; Gaps 19;

```
Qy 1 MSPNNEYEITLDASSSTVSVDNSVRYVPLANDQTTTLQNMNMYKDYLRMSGEENPELPGNP 60
Db 1 MTSRKNEEINLALSIPAVSNHSAQMLSTD-----ARIEDSLCTAEG-----NNI 47
Qy 61 BTPSSSTVQTGIGIVGQVLGALGVPFAGQIASPVSYFVIGQLMPSTSVSWEMIMKQVED 120
Db 48 DPVFSASTVQTGINIAGRILGVLGVPFAGQIASPVSYFVIGLWPRGR-DPWEIFLEHVQ 106
Qy 121 LIDQKITSVKRTALAGLOGLDGVYQKSLKWLNRNDTRASVVVTVQVIALELDFV 180
Db 107 LIQQVVTENTRTALARLQGLGNSFRAYQSLQEDMLENRDDARTSRVLYTVIALELDFL 166
Qy 181 AKISPAISQGVPLLSVYQAANLHLILLDASIFGAEMWFTPGCEISTFYDQVTRTAQ 240
Db 167 NAMPLPAIRNQEVPLLVYQAANLHLILLDASIFGSEFGLTSQEIORYYERQVEKTR 226
Qy 241 YSDYCVKWTGTLKGLKGTNAASWLVKQHPREMTLLVLDLVALFPNPDTRTYPIETTAQ 300
Db 227 YSDYCARWYNTGLANNLGRCTNAESWLVYQFPREDLTGLVLDLVALFPNPDTRTYPIETTAQ 286
Qy 301 LTRVYVTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSHPLFDILSBIEFYTTTAAQ 360
Db 287 LTRVYVTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSHPLFDILSBIEFYTTTAAQ 345
Qy 361 LPLANTVEYVWGHSHIKYKNTNASSALERNYGTITSNKIKYDVLANKDIFQVRSGLGADL 420
Db 346 -RWSNTQYMWYVGHSHIRSRIRGSLSTWTHGNTNTSINPVLTQTSRDVYRTSPAGIN 404
Qy 421 ANYTAQVGVYVYASFTLLDKNGTSGSGVGGFTYKSPHTMQVCTQNYNTIDIPPE--NEP 478
Db 405 ILLTPVNGVPWAFNWM--RNPLNSLRGSLLYTGYT--GVGTQLFDSLETPEPTTERP 460
Qy 479 LSRGYSHRLSHITSYSPSKNASSPARVGN---LVPFANTHRSADVNTVYSDKTIQIPV 535
Db 461 NYESYSHRLSNIRLIS-----GNTLRAPVYSWTHRSADRTNTISSDITQIPLV 509
Qy 536 KAHTLVSGTIVIKGPGFTGGMILKRTSSGPLAYTSVSVKSPLOKRYARIRYASTNMLRL 595
Db 510 KSFNLSGTSVSGPGFTGGDIIRTNVNGSVLSGMLNPNNTSLQRYRVRVRIASQTMVL 569
```

Qy	596	FVTLISGRIRIYSINVNKTNKGDDLTFTYTPOLATIGTAFTSPNTSYDSLTVAGDASFASGGEV	555
Dd	570	RVTVGSGTTTDOQPPTSMANESITLSQSFRFABFPVGISASG-SQTAGISISNNAQRQTF	628
Qy	656	VYDKFELLPYNATFEAREEDDVAKKANGLPFSKKO-ALQTSVTDYOVNQAAMLVECLSD	714
Dd	629	HFDKIEFIPIPTATUBAESDLERAKAVNALFTSNQIGLKTDVTHIDRVSNLVECLSD	688
Qy	715	ELYPNEXRMLWDVAKEARLUQARNLLQDTGFNRING--ENGWGTSGTIEVABGDVLFXD	772
Dd	689	EPCLDEKXELSEKVHKAHKLSDERNLLQDPNFRGINQLDRGWKSGTDTIQGGDDVPFKB	748
Qy	773	RSLELTSARELDHTETTYLYQQIDESLAPKYTKLGFIGSSODLEIKLIHRHANQIV	832
Dd	749	NYVTLGLTGD---ECYPITYLYQKIDESKLKAYTRYQLRGYIEDSQDLEYILHYRNARGHET	805
Qy	833	KNPVDNLLPOVLPVNSCGGIJDRCSQQYVDANLALENNGENGNMSSDSHAPSFIHDTEGI	892
Dd	806	VNVPGT--GSIMPLSASPICKCAHH-----SHIFSILDIDVGCT	842
Qy	893	DLNENTGIWVVVKLIPPTNGVATIGNLELVBEGLPSGSLTERAQQOOQWQDKMARKRGAAS	952
Dd	843	DLNEDLGWVVFIEKIKTDQGHARLGNLFEBEKLPGVGEALRAVRAEKRWDRKEKLEWET	902
Qy	953	EKAYYAQAQALDRFLADYQDQKLNBSGVEMSDMLAAQNLVOSIPVYNDALPELPGMYTTS	1012
Dd	903	NIVTYKAKESVDVALFVNSQYDRLOQDTNIAMIHADRKHVSIREAYLPELSVIPGVNAAI	962
Qy	1013	PTELTRLQAQAWNLYDLRENAIPNGDFPRLGSLDMWNATSDDNVV-QQLSDTSVLVIHPNNSQV	1071
Dd	963	PEELEGRIFATAFSLYDARNVINKGDFPNNGLSCMNWKGVHDVEEQNHRSVLVVPWEAEV	1022
Qy	1072	SQQTFVQPNTRYULRVITARKEGVDGYVIIIDGANQTFETLTFN-----ICDD	1118
Dd	1023	SQEVRVCPRGYILLURVAYKEGEGCVCTHEIENNTDELKFNSCIVEERYYPNNTVTCND	1082
Qy	1119	DTGV-----LSADQTS-----	1129
Dd	1083	YTATOBEYEGYTTSENRYGDAYESNSSVPADIYASAEKAYTDTGRDNPCBSNRGVGDY	1142
Qy	1130	-----YITKTVEFTPSTEQWMDMSSETGVFNIESVELWLBBE	1167
Dd	1143	TPLPAGYVTKBLEYFFPKDWIBIGISTEGTFVIDSVSELLIMEE	1186

Search completed: June 21, 2004, 10:23:47
Job time : 417 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 10:10:20 ; Search time 85 Seconds
(without alignments)
4331.879 Million cell updates/sec

Title: US-10-089-678-1
Perfect score: 6044
Sequence: 1 MSPNNQNEYILDASSSTVS.....MSBTEGVNIESVELVLEE 1167

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5903.5	97.7	1144	2	Q8KZL7
2	2750.5	45.5	1144	2	Q45745
3	2545.5	42.1	1228	2	Q33T75
4	2539.5	42.0	1228	2	Q33NM5
5	2433	40.3	1231	2	Q8KNY2
6	2056.5	34.0	1155	2	Q9F296
7	2050	33.9	1180	2	Q9S5V8
8	2048	33.9	1176	2	Q7WZT9
9	2034	33.7	1176	2	Q45736
10	2032.5	33.6	1155	2	Q9RC30
11	2032.5	33.4	1171	2	Q93T21
12	2016.5	33.4	1171	2	Q06894
13	2003	33.1	1169	2	Q8GHE8
14	1997	33.0	1189	2	Q3L877
15	1989.5	32.9	1174	2	Q45749
16	1953.5	32.3	1160	2	Q93TF9

17	1948.5	32.2	1128	2	Q9PDC0	Q9fde0 bacillus th
18	1939.5	32.1	1177	2	Q8GLY5	Q8g1y5 bacillus th
19	1937	32.0	1176	2	Q9S514	Q9s514 bacillus th
20	1933.5	32.0	1177	2	Q03743	Q03743 bacillus th
21	1933	32.0	1178	2	Q45768	Q45768 bacillus th
22	1931	31.9	1178	2	Q9R826	Q9r826 bacillus th
23	1930.5	31.9	1177	2	Q45735	Q45735 bacillus th
24	1925.5	31.9	1118	2	Q9AM83	Q9am83 bacillus th
25	1897.5	31.4	1118	2	Q9AM82	Q9am82 bacillus th
26	1836.5	30.4	1280	2	Q8VUK9	Q8vuk9 bacillus th
27	1780.5	29.5	1236	2	Q939T3	Q939t3 bacillus th
28	1746.5	28.9	1118	2	Q9AM81	Q9am81 bacillus th
29	1730.5	28.6	719	2	Q9F0P8	Q9f0p8 bacillus th
30	1727.5	28.6	719	2	Q85796	Q85796 bacillus th
31	1717.5	28.4	719	2	Q93NJ5	Q93nj5 bacillus th
32	1702.5	28.2	719	2	Q8KY61	Q8ky61 bacillus th
33	1702	28.2	1254	2	Q8VUL0	Q8vul0 bacillus th
34	1577.5	26.1	1270	2	Q8VUL1	Q8vul1 bacillus th
35	1378.5	22.8	638	2	Q87654	Q87654 bacillus th
36	1278.5	21.2	1155	2	Q9AM80	Q9am80 bacillus th
37	1254	20.7	645	2	Q9S603	Q9s603 bacillus th
38	1245	20.6	652	2	Q9S6N9	Q9s6n9 bacillus th
39	1171	19.4	723	2	Q9S4B5	Q9s4b5 bacillus th
40	1094	18.1	489	2	Q8KNV1	Q8knv1 bacillus th
41	1069.5	17.7	533	2	Q7X3F6	Q7x3f6 bacillus th
42	1061	17.6	381	2	Q45740	Q45740 bacillus th
43	1052.5	17.4	558	2	Q8VW63	Q8vw63 bacillus th
44	1042.5	17.2	526	2	Q32308	Q32308 bacillus th
45	1041	17.2	620	2	Q45720	Q45720 bacillus th

ALIGNMENTS

RESULT 1

Q8KZL7	PRELIMINARY;	PRT; 1144 AA.
ID	Q8KZL7	
AC	Q8KZL7;	
DT	01-OCT-2002 (TRENBLrel. 22, Created)	
DT	01-OCT-2002 (TRENBLrel. 22, Last sequence update)	
DT	01-OCT-2003 (TRENBLrel. 25, Last annotation update)	
DE	Cry8 protein.	
GN	Cry8.	
OS	Bacillus thuringiensis (subsp. galleriae).	
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.	
OX	NCBI_TaxID=29338;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=SDS-502;	
RA	Asano S., Yamamoto T.;	
RT	"a novel cry8 gene highly toxic to Anomala cuprea.";	
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AB089299; BAC07226.1; -	
DR	GO; GO:0015070; F:toxin activity; IEA.	
DR	GO; GO:0006952; P:defense response; IEA.	
DR	InterPro; IPR001178; Endotoxin.	
DR	InterPro; IPR005638; endotoxin C.	
DR	InterPro; IPR008979; Gal_bind_like.	
DR	Pfam; PF00555; endotoxin_1.	
DR	Pfam; PF03944; endotoxin_C_1.	
DR	Pfam; PF03945; endotoxin_N_1.	
SQ	SEQUENCE 1144 AA; 128060 MW; 98F93070C49014AB CRC64;	

Query Match 97.7%; Score 5903.5; DB 2; Length 1144;

Best Local Similarity 97.9%; Pred. No. 0;

Matches 1143; Conservative 0; Mismatches 1; Indels 23; Gaps 1;

Qy 1 MSPNNQNEYILDASSSTVSNDNRVRYPLANDQTTTLQNNMYKDYLRMSGENPBLFGNP 60

Db 1 MSPNNQNEYILDASSSTVSNDNRVRYPLANDQTTTLQNNMYKDYLRMSGENPBLFGNP 60

Qy 61 ETFISSSTVQTGIVGVQVIGALGVPFAGQIASFVSFVIGQLWPSSYSTVWEMIMKQVED 120

Db	61	ETPISSTVQTGIGIVGOVLGALGVPPAGQTASFYSFIVGQLPSPSTSVWEMIMKQVED	120
Qy	121	LIDQKITDSVRKTALAGLQGLGDLGVYQKSLKWLNRNDTRARSVVVYQYIALELDPV	180
Db	121	LIDQKITDSVRKTALAGLQGLGDLGVYQKSLKWLNRNDTRARSVVVYQYIALELDPV	180
Qy	181	AKIPSPAISSGOEVPLLSVYAAQANLHLLLRDASIFGAENGFTPGESTISTPYDQVTRTAQ	240
Db	181	AKIPSPAISSGOEVPLLSVYAAQANLHLLLRDASIFGAENGFTPGESTISTPYDQVTRTAQ	240
Qy	241	YSDYCVKWNVTGLDKLGKTNAASHLKTHQFRRENTLLVLDLVALFPNYDTRTPIETTAQ	300
Db	241	YSDYCVKWNVTGLDKLGKTNAASHLKTHQFRRENTLLVLDLVALFPNYDTRTPIETTAQ	300
Qy	301	LTRREVTDPPIVFNKRETSGGFCRRWSLNSDISFSVESAVIRSPHLPDILSEIEFPYTTFRAG	360
Db	301	LTRREVTDPPIVFNKRETSGGFCRRWSLNSDISFSVESAVIRSPHLPDILSEIEFPYTTFRAG	360
Qy	361	LPLNNTLEYVWGHISKYKXNTNASSALERNYGTITSNKIKYYDLANKDIFQVRSILGADL	420
Db	361	LPLNNTLEYVWGHISKYKXNTNASSALERNYGTITSNKIKYYDLANKDIFQVRSILGADL	420
Qy	421	ANYTAQVYGPYASFTLLDKNTGSGSVGGFTYSKPHITTMQCTQNYNTIDIEIPPENEPLS	480
Db	421	ANYTAQVYGPYASFTLLDKNTGSGSVGGFTYSKPHITTMQCTQNYNTIDIEIPPENEPLS	480
Qy	481	RGYSHRLSHITSYSPFSKNASSPARYGNLVPVAMTHRSADVTNTVYSDKITQIIPVVKAAHTL	540
Db	481	RGYSHRLSHITSYSPFSKNASSPARYGNLVPVAMTHRSADVTNTVYSDKITQIIPVVKAAHTL	540
Qy	541	VSGTTVIKGPFGTGNILKRTSSGPLAYTSVSKSPLSQRYRIRIRYASTTNLRLPVTIS	600
Db	541	VSGTTVIKGPFGTGNILKRTSSGPLAYTSVSKSPLSQRYRIRIRYASTTNLRLPVTIS	600
Qy	601	GTRIYSINVKNTMKGBDLPFTNTFDLATIGTAPTFPSNYSDSLTVGADSPASGGEVYVDKF	660
Db	601	GTRIYSINVKNTMKGBDLPFTNTFDLATIGTAPTFPSNYSDSLTVGADSPASGGEVYVDKF	660
Qy	661	ELIIPVNATFPAEEDLDVAKKAVNGLFTSKDALQTSVTDYQVNOAANLVCLSDSELYPNE	720
Db	661	ELIIPVNATFPAEEDLDVAKKAVNGLFTSKDALQTSVTDYQVNOAANLVCLSDSELYPNE	697
Qy	721	KRMLWDVAKKAKRLVQARNLLQDTGPNRINGENGTGSGLEVAEGDVLFKDRSLRLTSA	780
Db	698	KRMLWDVAKKAKRLVQARNLLQDTGPNRINGENGTGSGLEVAEGDVLFKDRSLRLTSA	757
Qy	781	REIDTETYPYLYQOIDESSLKPYTRYKLGKFIGSSQDLEIKLIRHRANOIVKXNPONLL	840
Db	758	REIDTETYPYLYQOIDESSLKPYTRYKLGKFIGSSQDLEIKLIRHRANOIVKXNPONLL	817
Qy	841	PDVLPVNSCGGIDRCSQOQYVDANLARNNGENGNMSDSSHAFSFIHDTGTEIDLNENTGI	900
Db	818	PDVLPVNSCGGIDRCSQOQYVDANLARNNGENGNMSDSSHAFSFIHDTGTEIDLNENTGI	877
Qy	901	WVVPFKIPTTNGYATLGNLVEBGLSGETLERAQOQEQQWQDKMKRKGASEKAYYAAK	960
Db	878	WVVPFKIPTTNGYATLGNLVEBGLSGETLERAQOQEQQWQDKMKRKGASEKAYYAAK	937
Qy	961	QAIDRLPADYQDQKLSNGVEMSDMLAQNVLVQSIPIPVYNDALPEIPGMNNTSFTELNRLL	1020
Db	938	QAIDRLPADYQDQKLSNGVEMSDMLAQNVLVQSIPIPVYNDALPEIPGMNNTSFTELNRLL	997
Qy	1021	QOAWNLVDLRANAIPNGFPRCLSDWNATSDVNVOOLSDTSVLVTPKNNSQVSOQFTVQPN	1080
Db	998	QOAWNLVDLRANAIPNGFPRCLSDWNATSDVNVOOLSDTSVLVTPKNNSQVSOQFTVQPN	1057
Qy	1081	YRYVLRVTARKEGVDGTVIIRDOGANOTETLTFNICDDDTGVLASADQTSVITKTVEPTPS	1140
Db	1058	YRYVLRVTARKEGVDGTVIIRDOGANOTETLTFNICDDDTGVLASADQTSVITKTVEPTPS	1117
Qy	1141	TEQWVIDKMSSETGVFNIESVELVLEE	1167

Db 1118 TEQWIDMSTEGVFNIESTVELVLEBS 1144

RESULT 2

QAS745 ID Q45745 PRELIMINARY; PRT; 1144 AA.

AC DT 01-NOV-1996 (TEMBLrel. 01, Created)

DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)

DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)

DB Delta-endotoxin (Pragmat).

GN CRYIX GENE.

OS Bacillus thuringiensis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1428;

RN [1]

SEQUENCE FROM N.A.

RA MEDLINE=94085596; PubMed=8262221;

RA Shestelev A.B., Svarinsky M.A., Karasin A.I., Kogan Y.N.,

RA Chetukhina G.G., Stepanov V.M.;

RT "Primary structure of the cryX--the novel Delta-endotoxin-related

RT gene from Bacillus thuringiensis ssp. galleriae.";

RL PDBS Lett. 336:79-82(1993).

DR EMBL; X75019; CAA52927.1; -

DR HSP; P07130; IDLC.

DR GO; GO:0015070; P:toxin activity; IEA.

DR GO; GO:0006952; P:defense response; IEA.

DR InterPro; IPR001178; Endotoxin.

DR InterPro; IPR005638; endotoxin_C.

DR InterPro; IPR005639; endotoxin_N.

DR InterPro; IPR008979; Gal_bind_Like.

DR Pfam; PF00555; endotoxin; I.

DR Pfam; PF03944; endotoxin_C; 1.

DR Pfam; PF03945; endotoxin_N; 1.

FT NON TER 1

SQ SEQUENCE 1144 AA; 129399 MW; 7D28594A19C4B065 CRC64;

Query Match 45.5%; Score 2750.5; DB 2; Length 1144;

Best Local Similarity 48.7%; Pred. No. 2.8e-148;

Matches 562; Conservative 197; Mismatches 341; Indels 53; Gaps 18;

Qy 41 NYKDYLRMSGEMPELFGNPETETISSVTGTGIGVQVLGALGVPPAGQIASFYSPFVG 100

Db 19 SYKDYLRKMSGDYIDSYINPQV--RTGLQTGIDIVAVVVGALGPGVGILTGLFLSLFG 76

Qy 101 QLPSPSTVSVMEMKQVEDLDQKITDSVRKTALAGLOGLGDLGVYKSLKQNLNRN 160

Db 77 FLFNSNDQAVWEAFIEQMEELEIQRISDQVVRTALDDLTGIGNYNYQYLALKWESERPN 136

Qy 161 DTRARSVVVQYTALELDLVAKTIPSPAISQ-----EVPLTSVYQAANLHLALLRDASI 215

Db 137 GVRA-NLVLRQRFILHALFVSSMPSFG-SGFSQRQFQQLLVVYQAANLHLALLADAEK 194

Qy 216 FGAEWGTFPGEISTFYDRQV--TRTAQVSDYCVKWYNTGLDKLXGTNAASWLKYHQPRRM 274

Db 195 YGARWGLRSEQIGNLYFNEILQTRDYTNHCVANNGLAGLGTSAESWLKYHQPREA 254

Qy 275 TLLVLVDLVALFPNYDRTYPIETTAQLTRVYTDPIPVNRETSGP-----CRKSLNSDI 330

Db 255 TLMAMDIALFPYNTRRYPPIANVPQQLTRVYTDPIPLGVPSSESLPEELCLRQWQTSAM 314

Qy 331 SPFEVSASVRSPLFDILSEIFPYTTRAGLPLNNTLEYLVWVGHSI--KYKQTNASSAL 388

Db 315 TFSLENALITSPHLPFTTNIMLYTGSFVHLTN-QLIEGWNIGHSVTSLLASGPTTVL 373

Qy 389 ERNYGTITSNKIYYDLANKDIPQVRS---LGLADLANYYAQVGVYPVASTLLDKNXGSG 445

Db 374 RRVNGSSTS-IVNYFSPNDRDVQJNTRSHTGLGPQN--APLFGITRAQF----- 420

Qy 446 SVGGFTYSKPHTTMQVCTQYNTIDEIP--PENEPISRGYSHRLSHITSVSF-----SK 497

Db 421 -YFGGTYVQTRNALTCQNYNSIDSLPDPNPEITSRSYSHRLSHITSYLRVLTIDGI 479


```
Db 340 MAGNAAPVL--VSLTGLGIFRTLSPLRYRIILGSGPNNQELFVLDGTBSPASLTNNL 397
Qy 443 GSGSVGGTYKPKHTTMOVCQNTYNTIDEIPENE--PLSRGYSHRLSHITSYSPSKNAS 500
Db 398 PS-----TIYRQRTV-----DSLVIIPQDNSVPPRAGFSHRLSHVTLMSQAAGAV 444
Qy 501 SPARYGNLPVPAWTHRSADVTNTVYSDKIQTIPVVKAHVTLVSGTVIKGPGFTGNILNR 560
Db 445 YTLR---APTFSQWHSRAEFNNIIPSSQITQITPLTKSTNLGSGTSVVRGPGFTGCDILR 501
Qy 561 TSSGPLATSVSKPSORSORRYRIRYASTNNLRFLVTSIGTRIVSINVNTKNGDDLT 620
Db 502 TSPQISTLRVNTIAPLSQRVRYRYASTNNLRFLVTSIGTRIVSINVNTKNGDDLT 561
Qy 621 FNTDLATIGTAFTPSNTSDSLTVCADSFASGGEVYVDKFLIPVATFEAEEDLDVAKK 680
Db 562 SGRFRTVCTPTFPNFSNGSVPTLSAHVFNSEVYIDRIEVPFAEYTFEAYDLERAQK 621
Qy 681 AVNGLFTSKD-ALQTSVTDYQVQAANLVECLSDLYPNKRMMDAVKAKRLVQARN 739
Db 622 AVNELFTSSNIGLTKTDVTHIDQVSNLVECLSDLYPNKRMMDAVKAKRLVQARN 681
Qy 740 LLODTGFNRING--ENGWTSSTGLVARGDVLKPDORSRLTSAREIDTETPTVLYQQID 797
Db 682 LLODPNFRGINRQDRGRGSDITIQGGDDVFKENYVTLGTFD---ECYPTLYQKID 738
Qy 798 ESKLPYRYKLGKFGIGSQDLLEIKLRANQIVKNVPDN----LLPDVLPVNSCGGID 853
Db 739 ESKLKAYTRYQLRGVIEDSQDLLEILRYNAKHETVNVPGTGLWPLSAQSPGKCGEPN 798
Qy 854 RSCSOQVVDANLALNENGNG-NNSSDSHAFSPHIDTGEIDLNTGNTGWVFKIPTNGY 912
Db 799 RCA--PHLEWNPDLDCSCRDGKCAHSHHPSLDIDVGCTDLNEDLGVVWVIFKIKTDGH 856
Qy 913 ATGLNLELVEEPLSGTLEAQQOEOQWODKMARKGKASEKAYAAQADRLPADYQD 972
Db 857 ARLGNEFLBEKPLVGEALARKAEKRWKREKLEMTIVYKBAKESVDALFVNSQY 916
Qy 973 QKLASGVMSQMLAQNLIQSPYVYNDALPEIPGMNTTSPETLNRQQAWNLYDLRAN 1032
Db 917 DOLQADTWIAMHAADKEVHSIRBAYPELSVIPGVNAAIFPELEGRIFTAFSLYDARNV 976
Qy 1033 IPNGDFRNLSDMNATSDNVV--QQLSDTSVLVFPNNSQVQOFTVQVNYRVLVATARK 1091
Db 977 IKNGDFNGLSCWNVKGVHDVVEEQNNQSRSLVVPWEAEVSOEVRVCPGCGYILVATYK 1036
Qy 1092 EGVGDGYVIRDGANQETLFPN-----ICDD-----1118
Db 1037 EGYGEGCVTHIEIENNTDELKPSNCVEEIVPNTVTNCNTVYNOBEYGGATSNRGNYN 1096
Qy 1119 DTGVLSDQTS-----YITKTVEFTPSTEQWID 1147
Db 1097 EAPSPADYASVYBEKSYDGERENPCFNRGDRYDTPLPVGYVTKELEYPEPDKWIB 1156
Qy 1148 MSETGVNIESVRLVLEEB 1167
Db 1157 IGETEGTFLVDSVELLAMEE 1176
```

RESULT 9

```
Q45736 ID Q45736 PRELIMINARY; PRT; 1176 AA.
AC Q45736;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Insecticidal crystal protein.
GN CRYIA(A).
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_Taxid=1428;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RA Kondo S., Tamura N., Kunitate A., Hattori M., Akashi A., Ohmori I.;
RT "Cloning and nucleotide sequencing of two insecticidal delta-endotoxin
RL genes from Bacillus thuringiensis var. Kurstaki HD-1 DNA.";
RN Agric. Biol. Chem. 51:455-463(1987).
RP SEQUENCE FROM N.A.
RA Udaasuuriyan V., Nakamura A., Mori H., Masaki H., Uozumi T.;
RT "Cloning of a new cryIA(a) gene from Bacillus thuringiensis strain
RL FU-2-7 and analysis of chimaeric CryIA(a) proteins for toxicity.";
RN Biosci. Biotechnol. Biochem. 58:830-835(1994).
DR EMBL; D17518; BAA04468.1; -.
DR PIR; JC2219; JC2219.
DR HSRP; P02965; 1CIY.
DR GO; GO:0015070; P:toxin activity; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; Endotoxin C.
DR InterPro; IPR005639; Endotoxin N.
DR InterPro; IPR008979; Gal bind_like.
DR Pfam; PF00555; endotoxin_1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
FT CONFLICT 77 77 L -> P (IN REF. 1).
FT CONFLICT 965 965 S -> F (IN REF. 1).
SQ SEQUENCE 1176 AA; 133075 MW; AA413594A289P85 CRC64;
```

Query Match 33.8%; Score 2045; DB 2; Length 1176;

Best Local Similarity 39.3%; Pred. No. 5.7e-108;

Matches 479; Conservative 174; Mismatches 415; Indels 152; Gaps 26;

```
Qy 53 NPGL-----FGNPET-FISSSTVQVG---IGI---VGQVLGALGVPPAGQIASFYSP 97
Db 4 NPNNECIPYCNLSNPVEVLGGERIEGYTPIDISLTQPLLEFVPGAGFVLGLVDI 63
Qy 98 IVGQLPSSVTSVWEMIMKQVEDLIDQKITSVRKNTALAGLQGLDGLDYQKSLKNWLE 157
Db 64 IWGIFGFSQ---WDAFLVQIEQLNQIRIBFARNQAISSLGLSNLYQIYAESFWEA 119
Qy 158 NENDTRASVVVTVQVIALELDPVAKIPSAISGQVPLSVVYAAANLHLRLDASIFG 217
Db 120 DPTPALRBERMRIQFNDMNSALTTAIPLLAVQNTVQVPLSVTVQANLHLSVLDRVSVEG 179
Qy 218 AEWGTFGEISTFYDRQVTRTAQSDYCVKQVNTGLDKLKGNTAAASLWKHQFREMILL 277
Db 180 QRWGDAATINSRYNDLTRLIGNYTDYAVRWYNTGLERWGPDSRDWRVYNQFRRELTLT 239
Qy 278 VLDVALFPNTDTRTYPIETTAQLTREVTDPVFNRETSGGCRRLSLNDSISFSEVE 336
Db 240 VLDVALFPNSDYSRRYPIRTVSQLTREIYTNVPLENFDGSPFGMAQRIEQN----- 290
Qy 337 SAVIRSPLHPLDLSIEEFT--TRAGLPLNNTLEYLVVGHSIK-----YKNTNASSALER 390
Db 291 ---IRQPHLMDILNSITITVDHVG-----FYMNGHQITASPVGSGEPFAPPLPG 339
Qy 391 NYGTTTSNKIKYDILANKDIFQVRS-----LGADLANYYAQQYGVYASFTLLDKNT 442
Db 340 MAGNAAPVL--VSLTGLGIFRTLSPLRYRIILGSGPNNQELFVLDGTBSPASLTNNL 397
Qy 443 GSGSVGGTYKPKHTTMOVCQNTYNTIDEIPENE--PLSRGYSHRLSHITSYSPSKNAS 500
Db 398 PS-----TIYRQRTV-----DSLVIIPQDNSVPPRAGFSHRLSHVTLMSQAAGAV 444
Qy 501 SPARYGNLPVPAWTHRSADVTNTVYSDKIQTIPVVKAHVTLVSGTVIKGPGFTGNILNR 560
Db 445 YTLR---APTFSQWHSRAEFNNIIPSSQITQITPLTKSTNLGSGTSVVRGPGFTGCDILR 501
Qy 561 TSSGPLATSVSKPSORSORRYRIRYASTNNLRFLVTSIGTRIVSINVNTKNGDDLT 620
Db 502 TSPQISTLRVNTIAPLSQRVRYRYASTNNLRFLVTSIGTRIVSINVNTKNGDDLT 561
```

```
Qy 621 FNTFDLATIGTAFPTFSNYSDSLTVGADSPASGGSEYVDKPELIPVNAATFAEEDLDVAKK 680
Db 562 SSGFRTVGTFTPFNFSGSVFTLSAHVFNSEYVYIDRIEFVPAEVTFFBAYDLERAQK 621
Qy 681 AVNGLFTSKKD-ALQTSVTDYQVNOAANLVECLSDLELYPNEKRMMDAVKEAKRLVQARN 739
Db 622 AVNELFTSSNQIGLKTDTVDYHIDQVSNLVECLDFECLQEKSEKVKHAKLSDERN 681
Qy 740 LLODTGPNRNG--ENGWTSSTGIEVAEGDVLKORSRLTSAREIDTETPTLYQQID 797
Db 682 LLODPNFRGINRQDRGWRGSTDTIQGGDDVFKNVYVTLTGTFD---ECYPTLYQKID 738
Qy 798 ESKLKPYTRYKLGKFGISSQDLLEIKLRHANOIVKQVDPN-----LLPDVLVFNSSCGGID 853
Db 739 ESKLKAYTRYQLRGYIEDSQDLLEYLYRYNAKHETVNVPGTSLWPLSAQSPIGKCGEPN 798
Qy 854 RCSEQQVVDANLALENGNG--NMSSDSHAFSPHIDTGEIDLNEMGTGIWVVFPIPTNGY 912
Db 799 RCA--PHLEWNPDLDCSRRGCKCNHSHHPSLIDVGCCTDLNEDLGWVIFPKIKTQDGH 856
Qy 913 ATGNLNLVEEGPLSGFTLERAQQOQWQDMARKGASEKAYAAQALDRLFADYQD 972
Db 857 ARGNLEFLSEKPLVGEALARKVRAEKWRDKREKLEWETNIVYKEAKESVDALFVNSQY 916
Qy 973 QKLSNGVMSDMLAAQNLVOSIPYVNDALPEIPCMNYTSTFELTNRLQOANLYDLRNA 1032
Db 917 DQLOADTNIAHIAADKRVHSIREAYLPBLSVPGVNAAI FEELEGRIFFAPSLYDARNV 976
Qy 1033 IPNGDFRNLGSDMNATSDVAV-QQLSDTVLVPNNNSQVQPTQPNRYVLRVATK 1091
Db 977 IKNGDFNGLSCNVKGVHVEEQNORSVLVPEWEAEVSQEVRCPCRGYILRVATYK 1036
Qy 1092 EGVGDGVYIIRDCANOTETLTFN-----ICDD-----1118
Db 1037 EGYGEGCVTHIEINNTDOLKFSNVEEBYVNNVTVCNDVTNQEYGGAYTSRNRGN 1096
Qy 1119 DTGVLSDAQT-----1119
Db 1097 EAPSVAPDAVSUYEEKSYTDGRRENPCFNRGVRDYTPLPVGVYVTKSLEYFPETDKVMIE 1156
Qy 1148 MSBETGVNIESVELVLEE 1167
Db 1157 IGTEGTFTVDSVELLLEE 1176
```

RESULT 10

```
Q9RC30
ID Q9RC30 PRELIMINARY; PRT: 1176 AA.
AC Q9RC30;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 135 kDa insecticidal protein.
OS Bacillus thuringiensis (subsp. kurstaki).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_Taxid:29339;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HD-1-02;
RA Hou B.K., Chen Z.H.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF154676; AAD55382.1; -.
DR HSSP; P02965; 1CIY.
DR GO; GO:0015070; P:toxin activity; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin C.
DR InterPro; IPR005639; endotoxin N.
DR InterPro; IPR008979; Gal bind like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
```

```
SQ SEQUENCE 1176 AA; 133010 MW; FCBEO69D0B81D8C4 CRC64;
Query Match 33.78; Score 2034; DB 2; Length 1176;
Best local Similarity 39.24; Pred. No. 2.4e-107;
Matches 478; Conservative 173; Mismatches 417; Indels 152; Gaps 26;
```

```
Qy 53 NPDL-----FQNPET-FISSSTVQTG---IGI-----VGQVLGALGVFPFAGQIASFYSF 97
Db 4 NPHNECIPYCNLSNPEVEVLGGRIETGTPIDISLSLTQFLSEFVGAGFVLGLVDI 63
Qy 98 IVQGLWPSSTVSVMEMIMKQVEDLIDQKIDTSVRKTKALAGLQGLDGLDVTQKSLNMLE 157
Db 64 IWGIFGSPQ---WDTFLVQIEQLINQIRIBEFARNAQISRLGLESLNYQIYAESPREMEA 119
Qy 158 NRNDTRARSVVVQYIALELDFVAKIPSAISGOEVLPSVYQAANLHLLALDRDASIFG 217
Db 120 DPTNPALREKRIQFQDNMSALTTAIPLLAVQNVQVPLLSVYQAANLHLSVLRDVSVFG 179
Qy 218 AEWGFTPEISTPYDRQVTRTAQYSDYCVKMYNTGLDKLKGNTAASMLKTHQFREMTELL 277
Db 180 QRWGFDAAATINSRNDLTRLIGNYTDYAVRWYNTGLERWGPDSRDWRYNQFRELIT 239
Qy 278 VLDELVALFPNYDTRTYPIETTAQLTREYTDPIVFNRETS--GGFCRRWSLNSDISFSBVE 336
Db 240 VLDIVALFSNVDSRRYPRTVSQLTREIYNPVLNFDGSPRGMARIEQN-----290
Qy 337 SAVIRSHLFDILSEIEFYT-TRAGLPLNNTLEYLVWGHSHIK-----YKNTNASSALER 390
Db 291 ---IROPHLMDILNSITITVDVHRG-----PNYMSGHQITASPVGFSQPEFAPLFG 339
Qy 391 NYGHTITSNKLKYDLANKDIFQVRS-----LGADLANYYAOVYGVVPYASFLLDKNT 442
Db 340 MAGNAAPPVL--VSLTGLGIFRFLSSPLRYRILLCGSPNNQELPVLDTETESFASLTN 397
Qy 443 GSGSVGFTYSKPHNTMQVCTQNTYNTIDEIPPENE--PLSRGYSHRLSHITSYSFKNAS 500
Db 398 PS-----TIYRQGTV-----DSLDTVIPPQDNSVPPRAGFSHRLGHVMTLSQAAGAV 444
Qy 501 SPARYGNLPVPAWTHRGADVTNTVYSDKIQTQIPVYKAHTLVSGTIVIKGPGFTGNILKR 560
Db 445 YTLR---APTFSWQHSABFNNIIPSSQITQIPTKSTNLGSGTSVVRKPGFTGGDIILRR 501
Qy 561 TSSGPLAYTSVSVKSPLSQRYRARIYASTTNLRLFVTISGTRIYSINVNNTMKGDDLT 620
Db 502 TSPQGISLIRVNIITAPLSQRYRARIYASTTNLRFHTSIDGRPINQGNFSAATMSSGSNLQ 561
Qy 621 FNTFDLATIGTAFPTFSNYSDSLTVGADSPASGGSEYVDKPELIPVNAATFAEEDLDVAKK 680
Db 562 SSGFRTVGTFTPFNFSGSVFTLSAHVFNSEYVYIDRIEFVPAEVTFFBAYDLERAQK 621
Qy 681 AVNGLFTSKKD-ALQTSVTDYQVNOAANLVECLSDLELYPNEKRMMDAVKEAKRLVQARN 739
Db 622 AVNELFTSSNQIGLKTDTVDYHIDQVSNLVECLDFECLQEKSEKVKHAKLSDERN 681
Qy 740 LLODTGPNRNG--ENGWTSSTGIEVAEGDVLKORSRLTSAREIDTETPTLYQQID 797
Db 682 LLODPNFRGINRQDRGWRGSTDTIQGGDDVFKNVYVTLTGTFD---ECYPTLYQKID 738
Qy 798 ESKLKPYTRYKLGKFGISSQDLLEIKLRHANOIVKQVDPN-----LLPDVLVFNSSCGGID 853
Db 739 ESKLKAYTRYQLRGYIEDSQDLLEYLYRYNAKHETVNVPGTSLWPLSAQSPIGKCGEPN 798
Qy 854 RCSEQQVVDANLALENGNG--NMSSDSHAFSPHIDTGEIDLNEMGTGIWVVFPIPTNGY 912
Db 799 RCA--PHLEWNPDLDCSRRGCKCNHSHHPSLIDVGCCTDLNEDLGWVIFPKIKTQDGH 856
Qy 913 ATGNLNLVEEGPLSGFTLERAQQOQWQDMARKGASEKAYAAQALDRLFADYQD 972
Db 857 ARGNLEFLSEKPLVGEALARKVRAEKWRDKREKLEWETNIVYKEAKESVDALFVNSQY 916
Qy 973 QKLSNGVMSDMLAAQNLVOSIPYVNDALPEIPCMNYTSTFELTNRLQOANLYDLRNA 1032
Db 917 DQLOADTNIAHIAADKRVHSIREAYLPBLSVPGVNAAI FEELEGRIFFAPSLYDARNV 976
```

```

Qy 1033 IPNGDFRGLSDMNATSDVNV-QQLSDTSVLVIFPNWNSQVQFTVQPNRYVILVLTARK 1091
Db 977 IKNQDFNGLSCWNVKGHVDVEBQNNQSRVLVVPWEAEVSEVRCVCGYILRVTAJK 1036
Qy 1092 EGVGUGGVIIIDGANTOTETLTPN-----ICDD----- 1118
Db 1037 EGYEGCVTHIEINNTDLKFSNCVBEIYSSNNVTTCNDYTVNOBEYGGAYTSRNRGYN 1096
Qy 1119 DTGVLSDADOTS-----YITKTVETFPSTEQWID 1147
Db 1097 EAPSPADYASVYEKSYTDCRRNPCEFRNGRYDTPLPVGYVTKLEYEPETDKWIE 1156
Qy 1148 MSETGVFNIESVELVEE 1167
Db 1157 IGETGTFIVDSVELLME 1176

RESULT 11
Q93T21 PRELIMINARY; PRT; 1155 AA.
AC Q93T21;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Crystal protein CryIab16.
GN CryIab16.
OS Bacillus thuringiensis (subsp. israelensis).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1430;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=407;
RA Yu J., Tan L., Wu D., Pang Y.;
RT "Molecular characterization of a silent gene encoding a 130-kilodalton
RT crystal protein from Bacillus thuringiensis subsp. israelensis.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF375608; AAK55546.1; -.
DR PIR; A29125; A29125.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin C.
DR InterPro; IPR005639; endotoxin N.
DR InterPro; IPR008979; Gal bind_like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin C; 1.
DR Pfam; PF03945; endotoxin N; 1.
SQ SEQUENCE 1155 AA; 130747 MW; 7FC98E0100C7698 CRC64;

Query Match 33.6%; Score 2032.5; DB 2; Length 1155;
Best Local Similarity 38.9%; Pred. No. 2.9e-107;
Matches 473; Conservative 185; Mismatches 392; Indels 167; Gaps 29;

Qy 53 NPGL-----PGNPET-FTSSSTVQVG---IGI---VGQVLGALGVPPAGQIASFVSP 97
Db 4 NPNTNECIPYNCLSPNEVEVLGGERIETGYTPIDISLQTPLLSEFPVPGAGFVLGLVDI 63
Qy 98 IVGQLWPSSTSVSWKMKQVBDLIDQKITDSVRKATAGLQGLGDLVDVQKSLKNWLE 157
Db 64 IWGIFGPQQ----WDAFLVQIQLNQRIEERFARNQASRLGLESLNLYQIYAESFWEA 119
Qy 158 NRNDTRARSVVVQTYIALELDFVAKIPSPAISQGVPLSVYAQAANHLHLRLDASIFG 217
Db 120 DPTNPALREEMRIQPNDMNSALTATPLFAVQNYRVLPLSVYVQAVNHLSLVDVLVFG 179
Qy 218 AEWGTFPGESTIFTRQVQTRTAQYSDYCVKWTNTGLDKLKTNAASWLKYHQFREMILL 277
Db 180 QRWGFDAAATINSRYNDLRLIGNYTDHAWRYNTGLERVMGDPDSRDRIYNOFREELTLT 239
Qy 278 VLDLVALPNTYTRTYPIETTAQLREVVTDPVFNRTSGGFCRRWSLNSDISFSVEBS 337
Db 240 VLDIVSLFPNYSRTYPIRTVSQLTREIYNPLENFD--GSF--RGSAGQ-----IEG 289

```

```

Qy 338 AVIRSPHLFDILSBIFBTTRAGLPLNNTLEYVWGHSHK-----YKNTNASSALERY 392
Db 290 S-IRSPHMDILNSITTYT-----DAHRGEY---YMSGHQIMASPVGFSGPEFTFPL---Y 338
Qy 393 GTIISNKIKYVDLANKDIPVRSILGADLANYYAQVY--GVVPASFTLLDKMT-----GSGS 446
Db 339 GTMGNAAPQORIVAQLOGGVTRTISSTL---YRPFNIGINNQQLSVLDTGTFAYGTSN 395
Qy 447 VGGFTYSPHTTMQVCTQNTYNTIDIBPPENB--PLSRGYSHRLSHITSY--SPSKNASSP 502
Db 396 LPSAVYRKSGTV-----DSLDEIPQNNNVPPROGFSHRLSHVSMFRSGFSNVS61 447
Qy 503 ARYGNLVPFAWTHRSADVNTVYSDKIQIPVKAHTLVSGTIVTKGPGTCGNILKRTS 562
Db 448 IR---APMFSWIHRSAEFNNIIPSSQITQIPLTKSTNLGSGTSVVKGPGTGGDILRRTS 504
Qy 563 SGPLAYTSVSVKSPLSQRYRIRYASTTNLRLFTVITSGTRIYSINVNKNKMGDDLTEN 622
Db 505 PQQISTLAVNITAPLSQRYRIRYASTTNLQFHTSIDGRPINQGNPFSATMSGSLQSG 564
Qy 623 TFDLATIGTFTFSPNSDSLTVGADSPASGGEVYVDKPELIPVNATPFAEBOLDVAKAV 682
Db 565 SFRTVGFTTFFNFSGSVFTLSAHVPNSGNEVYIDRIEFVPAEVTFEABYDLERAQAKV 624
Qy 683 NGLFTSKED-ALQTSVTDYQVNOAANLYECLSDLEYLPHKEKMLWDVAKELVQARNLL 741
Db 625 NELFTSSNQIGLKTVDVTDYHIDQVSNLYECLSDLEFCLDEKCKLSKVKHAKLSDRNLL 684
Qy 742 QDTGFNRING--ENGWGTSGTIEVAEGDVLKDRSLRLTSAREIDTETPTLYQQIDES 799
Db 685 QDPNFRGINQLDRGWRGSTDITIOGGDDVPKENTVTLTGTPD---ECYPTLYQKIDES 741
Qy 800 LLKPYTRYKLKGFIGSSQDLKELIRHRANQIVKRVNPNLLPDLVFNWNSCGGIDRCSEBQ 859
Db 742 KLKATRYQLRGYIBDSQDLRIYLIRYNAGHETVAVPGT--GSLWPLSAPSPIGKCAHH- 798
Qy 860 YVDANLALENNGENGMSDSSHAPSFHIDTGEIDLANENTGIWVVPKIPPTNGVATLGNLE 919
Db 799 -----SHHFLSDIDVCGCTDLNEDLGWVVIPIKIKTODGCHARLGNLE 838
Qy 920 LVEEGPLSGETLERAQOQBOODKMARKGASERKAYVAAKQAIIDRLFADYDQKLSGV 979
Db 839 FLEEKPLVGEALARVKRAEKWRDKREKLEWTNIVYKEAKESVDVAVFVNSQYDLQADT 898
Qy 980 EMSMDLAAQNLVQSIPTVYVNDALPBIPGMYNTSFTTELTRLQOQANLYDLRNAIPNGDPR 1039
Db 899 NIAMIHAADKRVHSIRBAYLPBLSVPIGVNAAIPELEGRITPAPSLYDARNVIRKNGDEN 958
Qy 1040 NGLSDWNATSDVNV-QQLSDTSVLVIFPNWNSQVQFTVQPNRYVILVLTARKEGVDGY 1098
Db 959 NGLSCWNVKGRHVDVEEQNNHRSVLVVPWEAEVSEVRCVCGYILRVTAKEYGEGC 1018
Qy 1099 VIIRDGAMOTETLTEN-----ICDDDTGV----- 1122
Db 1019 VTIHRIENYAELEKFCNVEEVEVPNNVTVCNDYATQOEYEGTYTSRNRGYDGAYESNS 1078
Qy 1123 -LSADOTS-----YITKTVETFPSTEQWIDMSE 1150
Db 1079 SVPADYASAEKAYTDCRRNPCEBNSRNGYDGYTLPAGYVTRLEYFPETDKWIEIGE 1138
Qy 1151 TEGVFNIESVELVEE 1167
Db 1139 TEGTPIVDSVELLME 1155

RESULT 12
Q06894 PRELIMINARY; PRT; 1171 AA.
AC Q06894;
DT 01-JUL-1997 (TremBLrel. 04, Created)
DT 01-JUL-1997 (TremBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)

```


305 NNITFTDFSVGRN-----FTMGHHRVSSLIIGGNNITSPYGRZANQBPFRSFTENG 359
410 IFQVRSIGADLANYYAQQVGVYASFTLLDQWGTSGVGGFTYSGKPHHTMQVCTQNTI 469
360 VP--RILSNFTLRLQWPAAPPNLGVBGVFSTPTNSFTYRGRGTV-----DSL 409
470 DEIPPENE--PLSRGYSHRLSHITSYSFKNASSPARYGNLPVFAWTHRSADVNTVYSD 527
410 TELPPEDNSVPREGYSRLLCHAT---FVQRSGTPTLTGCV-VFSWTHRSATLTNTIDPE 465
528 KTOIPVKAHTLVSTTVIKGGFTGNTLKRTSSGFLAYTSVSKSPLSQRYRARI 587
466 RINQIPLVKGFVRWGGTSTVTCGFTGGDILRNTFGDFVSLQVNSPITQRYLRFY 525
588 ASTNLRLFTV-----ISGTRIYSINVKWTKNGDDLTFTNTFDLATIGATTPSNYS 640
526 ASRDARVIVLGAASGTGQGVSVNMPLOKTEIGENLTSRTFTYDFSNPFSFRANPD 585
641 SLTV-----GADSPAGSGEVYVDKPELIPVNAATFAEEDLDVAKAVNGLFTSKD-AL 693
586 IIGISERPLFGAGSISS-GELYDKIEIILADATFAESDLERAQKAVNALFTSSNOIGL 644
694 QTSVTDYQVNOANLVCLSDSELYPNEKMLMDAVKEAKRLVOARNLLQDTGFNRIN--G 751
645 KYDVTYDHIDQVSNLVKLSDEFCLDEKRELSKVKHAKRLSDERNLLQDPNFRGINRQP 704
752 ENGTWGTGIEVAEGDVLKORSRLTSABRIDETPTVLYQIDESLLKPTRYKLKG 811
705 DRGWRGSTUITTQGGDDVFKENVTUPTGT--VD-ECYPTYLYQKIDESLLKATRYELRG 761
812 FTGSSQDLBKILIRHRANOIVKNVDPN-----LIPDLVLPVNSCGGIDRCSEQQVVDANLAL 867
762 YTEDSQDLBIYLIRYNAKHEIVNPGTSLMPLSAQSPGKCGEPNRCA--PHELVNPD 819
868 ENNGENG-NWSSDSHAFSHIDTGETIDLNENTGIWVFKIPTTNGYATLGNLELVEEGPL 926
820 DCSRDGKCAHSHHFTLIDVCGTDLNEDLGWVIFKIQDGHARGNLNLEFLKEXPL 879
927 SGTETLERAQQOQWQDKWARKGSEKAYYAAQKQIDRLPADYQDKLNSGVHESDMLA 986
880 LGEALARVRAEKWRDKREKLOLETNIVYKAKESVDALFVNSQYDRQLQVDTNAMIHA 939
987 AQNLVQSIPYVNDALPRIPKANYTSFTLTVRLAQAMNLYDLRNAIPNGDFRNLSDWN 1046
940 ADKRVHIREALPELSVTPGVNAAIFPELEGRIFTAYSLYDARVINKGDFRNLGCLWN 999
1047 ATSDVNV--QQLSDTSVLVLPNNWSQOQFTVQPNRYVLRVTARKEGVGDGWIIRDCG 1105
1000 VKGHVDVEQNNHRSVLVPEWEAEVSQVRCPCRGVILRTVTAKEGVGECVTIHEIE 1059
1106 NOTETLTN-----ICDDDTGV-----LSADOT 1128
1060 DNTDELKFSNVEEVEVPNNVTTCNNYTGTOBEYEGTYSRRNQGYDEAFGNPNPVPADYA 1119
1129 S-----YITKVEFTPTSTROVWIDMSETGEVNI 1157
1120 SVYEKSYTDGRRNPNCSNRYGDTPLPAGYVTKDLFTPTDKWIBIGETEGTPIV 1179
1158 EVELVLEER 1167
1180 DSVBLVLEE 1189

RESULT 15
Q45749
ID Q45749 PRELIMINARY; PRT; 1174 AA.
AC Q45749;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Crystal protein (Crystal delta-endotoxin).
GN CRY1PB.
OS *Bacillus thuringiensis*.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BTS00349A;
RA Lambert B.;
RT "NO INFORMATION";
RL Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=B-Pr-88;
RA Li C., Zhang J., Huang D., Li G.;
RT "A crystal endotoxin from *Bacillus thuringiensis* strain B-Pr-88.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z22512; CA80235.1; -;
DR EMBL; AF336114; AA013295.1; -;
DR PIR; S32649; S32649.
DR HSP; P02965; ICIY.
DR GO; GO:001070; P:toxin activity; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; Endotoxin C.
DR InterPro; IPR005639; Endotoxin N.
DR InterPro; IPR008979; Gal_bind_Like.
DR Pfam; PF00555; Endotoxin; 1.
DR Pfam; PF03944; Endotoxin C; 1.
DR Pfam; PF03945; Endotoxin N; 1.
SQ SEQUENCE 1174 AA; 13350 MW; 8C7F122F9446F15C CRC64;

Query Match 32.9%; Score 1989.5; DB 2; Length 1174;
Best Local Similarity 38.0%; Pred. No. 8.4e-105;
Matches 458; Conservative 191; Mismatches 413; Indels 143; Gaps 25;

QY 59 NPETFISSSTVQTG-----IGIVQVLGALGVPPAGQIASFYSFTVQGLMPSTVSW 111
DB 17 NPEVELLSERSTGRPLDISLTFLLSRP-VPGVGVAFLDLMGFTTPE-----W 71
QY 112 ENIMKQVEDLIDQKITSVRKTALAGLQGLDGLVYQKSLKNLKNRNDTRARSVVVQT 171
DB 72 SLFLQLQIEQLQRIETLERNRAITTLGLADSYEVYLEALREWEENPNNAQLREDVRIR 131
QY 172 YLALSLDPVAKIPSPAISQVFPILLSVYQAANLHLLLRDASIFGAEMGFTTPEISTFY 231
DB 132 FANTDADALITAINNFTLSFBIPLLSVYVQAANLHLLLRDASIFGAEMGFTTPEISTFY 191
QY 232 DRQVTRTAQSDYCVKNTVGLDKLGTNAASWLKXKHOFRRMTLAVLVALFPNYDTR 291
DB 192 NRLNLIHRYEHCDDTNOGLENLKGTNTROWSPNQPRRLTLTLVLDIVALLFPNYDAR 251
QY 292 TYPIETTAQLTREVTYDPIVFNRETSGGFCRWLSNLSDISPSESVSAVIRSPLDILSE 351
DB 252 AYPITQSQTREIYVTSVIEDSPVSA-----NIPNGFNRAEFGV-RPPLADPMNS 302
QY 352 IEFVTRAGLPLANNTEYLEVWGHISIKYNTVASSALERNYGTIT-----SNKIKY 402
DB 303 L-FVTA-----ETVRSQTVMGGLVSSRNAGNPINFPYIPNGGAIWADHPRPP 355
QY 403 YDLANKOIFQVRSIGADLANYYAQQVGVYASFTLLDQWGTSGVGGFTYSGKPHHTMQV 462
DB 356 YRTLSDPVF-VRG-GFGNPHVTLGRGVAF-----QQTGTN-----HTRTPRN 396
QY 463 TONYNTIDEIPPENEPLS--RGYSIRLSHITSYSFKNASSPARYGNLPVFAWTHRSADV 520
DB 397 SGTIDSLDEIPPDNSGAPNDYSHVLNHTVFWRPGBIAGSDSN-RAPMFWTHRSADR 455
QY 521 TNYVYSKTIQIPVVKAAHTLVSTTVIKGGFTGNTLKRTSSGFLAYTSVSKSPLSQRY 580
DB 456 TNIINPNIIQIPAVYKAHTLVSTTVIKGGFTGNTLKRTSSGFLAYTSVSKSPLSQRY 515
QY 581 YRARIYASTTNLRLFTVITISGTRIYSINVKWTKNGDDLTFTNTFDLATIGATTPSNYS 640
DB 516 YRVRIYASTTDLQFPFTRINGTSVNOGNFQRTMRNGNLESGNFTAGTFTPTPSFSAQS 575

```
Qy 641 SLTVGADSPASGGGVYUDKFLIPVNFATFEAEEDLDVAKAVNGLPTSKD-ALQTSVTD 699
Db 576 TFLTGTOAP-SNOGVYIDRIEFVPAEVTFEAESDLERAQKAVNALPTSTSQLGKTNVTG 634
Qy 700 YQVQAANLVECLSDLYPNKRMWDVAKELVQARNLLQDTGFNRIN--GENGWVG 757
Db 635 YHIDQVSNLVACLSDPCLDEKRLSKVKHAKLSDKGNLLQDPNFRGINRQPDHGWRG 694
Qy 758 STGIEVABGVLPKDRSLRLTSAREIDTETVPTYLYQOIDESLLKPYTRYKLGKFISSQ 817
Db 695 STDITIQGGDDVPKENVYTLPGTFD--ECYPTLYQKIDESKLKAYTRYQLRGYIEDSQ 751
Qy 818 DLEIKLIRHRANQIVKNVP--DNLLPDVL--PVNSCGIDRCSEQOQYVDANLALENNEN 873
Db 752 DLEIYLIRYNSKHIVNVPVGTGSLWPLSVENQIGPCGEPNRCA--PHLEWNPDLHCSCR 809
Qy 874 GNMS-SDSHAPSPHIDTGEIDLNENTGIWVVPKIPTTNGVATLGNLELVBEGLSGETLE 932
Db 810 GEKCVHSHHPSLDIDVGCTDLNEDLGWLLIFKIKTQDGHARLGNLBEFLBEEPLGHALA 869
Qy 933 RAQOQEQWQDMARKGASEKAYYAAQAIIDRLFADYDQDKLNSGVMSDMLAAQNLVQ 992
Db 870 RVKEAEKWRDKREKLQLETNIVYKEAKESVDALFVNSQYDLQADTNMIAMHAADKRVH 929
Qy 993 SIPIVYNDALPEIPGMNYTSTFELTNLQOAWNLYDLNNAIPNGDFRNLSDNATSDVN 1052
Db 930 RIREAYLPFLSVIPGVNAAIPEELEGRIFTAYSIDARNVKNGNFNNGLLCNVKGHRVD 989
Qy 1053 V-QQLSDTSVLVFNKNSQVQFTVOPNRYVLRVTARKGVGDGVYIIRDGANQTTTL 1111
Db 990 VEQNNHRSVLVPEWEAEVSQKRVCPGRGYILLRVAYKEGYGEGCVTIHEISDNTDEL 1049
Qy 1112 TF----- 1113
Db 1050 KFSNCVEEGYPNNTVTCEYTMNQVGCCTDACNVRNRYEDAYGHNPTPVHYTTPVEE 1109
Qy 1114 -----NICDDDTGVLSAD--QTSYITKTVFTPTSTEQVWIDMSBTEGVFNIESVEL 1162
Db 1110 EYTDERRNPCEANKGVNVTPLFVGVTYKLEYPPETDTVMIEIGETGTFIVDSVEL 1169
Qy 1163 VLEEE 1167
Db 1170 LLMEE 1174
```

Search completed: June 21, 2004, 10:12:34
Job time : 94 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:35:05 ; Search time 28 Seconds
(without alignments)
4009.125 Million cell updates/sec

Title: US-10-089-678-1
Perfect score: 6044
Sequence: 1 MSPNNQNEYIILDASSTSV.....MSETGVFNIESVELVLEE 1167

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues 283366
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3733.5	61.8	1160	2 I40589	parasporal crystal
2	3123	51.7	1157	1 S49247	parasporal crystal
3	2750.5	45.5	1154	2 S39536	parasporal crystal
4	2547.5	42.1	1228	2 S00873	parasporal crystal
5	2458.5	40.7	1138	2 A48944	parasporal crystal
6	2191	36.3	1156	2 S19306	parasporal crystal
7	2154.5	35.6	1174	2 A42459	parasporal crystal
8	2093.5	34.6	1166	2 S32645	parasporal crystal
9	2064.5	34.2	1155	2 A26513	parasporal crystal
10	2061.5	34.1	1155	2 JD0002	parasporal crystal
11	2052.5	34.0	1155	2 S02134	parasporal crystal
12	2048	33.9	1176	2 JT0241	parasporal crystal
13	2045	33.8	1176	2 JC2219	parasporal crystal
14	2044.5	33.8	1181	2 A41052	parasporal crystal
15	2043	33.8	1156	2 A29125	parasporal crystal
16	2038	33.7	1176	2 S02215	parasporal crystal
17	2037	33.7	1176	2 A22617	parasporal crystal
18	2030.5	33.6	1155	2 I39838	parasporal crystal
19	2024.5	33.5	1171	2 I40572	parasporal crystal
20	2010.5	33.3	1171	2 A37829	parasporal crystal
21	1989.5	32.9	1174	2 S32649	parasporal crystal
22	1984	32.8	1189	2 S00944	parasporal crystal
23	1953.5	32.3	1160	2 S32647	parasporal crystal
24	1944	32.2	1165	2 S11446	parasporal crystal
25	1939	32.1	1178	1 USBSKH	parasporal crystal
26	1932	32.0	1156	2 A29838	parasporal crystal
27	1930.5	31.9	1177	2 A49785	parasporal crystal
28	1923.5	31.8	1176	2 A48970	parasporal crystal
29	1917	31.7	1172	2 S32689	parasporal crystal

30 1868.5 30.9 1136 1 USBS81 parasporal crystal
31 1778.5 29.4 1180 2 I39870 parasporal crystal
32 1772.5 29.3 1180 2 A26858 parasporal crystal
33 1734.5 28.7 719 2 I40590 cryV465 protein -
34 1722.5 28.5 719 2 I39815 insecticidal prote
35 1713.5 28.4 719 2 S25383 parasporal crystal
36 1707.5 28.3 719 2 I39814 insecticidal prote
37 1683.5 27.9 934 2 A22798 parasporal crystal
38 1413 23.4 823 2 S04181 parasporal crystal
39 1372.5 21.1 652 2 I39811 parasporal crystal
40 1258.5 20.8 659 2 S10228 parasporal crystal
41 1245 20.6 652 2 A27323 parasporal crystal
42 1132.5 18.7 934 2 B29838 parasporal crystal
43 1116.5 18.5 649 1 JH0261 parasporal crystal
44 1069.5 17.7 655 2 JC7140 protoxin - Bacillu
45 1061 17.6 380 2 B42459 hypothetical prote

ALIGNMENTS

RESULT 1

I40589
parasporal crystal protein cry8Cal - Bacillus thuringiensis
N;Alternate names: parasporal crystal protein cryIII
C;Species: Bacillus thuringiensis
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 01-Dec-2000
C;Accession: I40589
R;Sato, R.; Takeuchi, K.; Ogiwara, K.; Minami, M.; Kaji, Y.; Suzuki, N.; Hori, H.; Asan
Curr. Microbiol. 28, 15-19, 1994
A;Title: Cloning, heterologous expression, and localization of a novel crystal protein
A;Reference number: I40589; MUID:94100786; PMID:7764305
A;Accession: I40589
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1160 <RES>
A;Cross-references: EMBL:U04366; NID:G532523; PIDN:AAA21119.1; PID:G532524
C;Superfamily: parasporal crystal protein
C;Keywords: delta-endotoxin

Query Match 61.8%; Score 3733.5; DB 2; Length 1160;
Best Local Similarity 64.2%; Pred. No. 1.6e-208;
Matches 767; Conservative 109; Mismatches 257; Indels 61; Gaps 21;

QY 1 MSPNNQNEYIILDASSTSVSDNSRYPLANDQTTLLQNNYKDYLRMSGEENPELPGNP 60
DB 1 MSPNNQNEYIILDASSTSVSDNSRYPLANDQTTLLQNNYKDYLRMSGEENPELPGNP 60
QY 61 ETPISS-STVQTGIGIVGVGLGALGVPPAGQIASFYFIVGQLWPSSTVSWEIMKQVE 119
DB 61 GTPISSAQDAVGTGIDIVSTIIISGLGIPVLGEVFSILGSLGLLWPSNNENWQIPMKRVE 120
QY 120 DLIDQKITTSVRKTLAAGLQGLDGVYQKSLQWLENRNDTRARSVVVQYIALELDF 179
DB 121 ELIDQKITLSVRGRATADLANSRIAYEYQNALEDKFNPHSTSAALVKRFGNABAIL 180
QY 180 VAKIPSPAISGQSVPLLSVYAQAANLHLLLRDASIFGAEWGFTPGGISTFYDQVYTTA 239
DB 181 RTNMGSEFSQNTYETPLPTVAQAASLHLLVMDVQIVGKEWGPQNDIDLFEKQSVYTA 240
QY 240 QYSDYCVKQNTYGLDKLKTNAASWLKHQFRENTHLLVDLVALFNYDTRTYPIETTA 299
DB 241 RYSDHCQVQWYAGNAGLKLGRGQWVDYRFRNEMVMVLDLVALFNYDARIIPLETNA 300
QY 300 QLTREYVYTDPIVFNRETSGGFCRRWSLW----SDI-----SFESEVSARSIRPHLFDIL 349
DB 301 ELTREYVYTDPIVFNRETSGGFCRRWSLW----SDI-----SFESEVSARSIRPHLFDIL 352
QY 350 SEIEFTYTRAGLPLNNT-EYLEYVWGVHSIKYKNTNASSALERNGYTTISNKIKYDILANK 408
DB 353 QEIRMYTS---PRQMGTYIEYNYWGGQRLTSLYVYGSSP--NKYSGLVAGAEIDIIPVQCN 407
QY 409 DIFQVRSGLADLANYYAQVGVYPASFTLLDKVGTSGSGVGGFTYSKPHHTTQVCTQNYNT 468

Db 408 DIYRV--VMTYIGRYTMSLLGVNVPVT- YFSNNTKQ-----TYSKP-----KQFAGGIKT 454

Qy 469 ID---EIPPENEPLSGVSHRLSHITSYSFSKNASSPARYNLPVAWTHRSADVNTTVY 525

Db 455 IDSGEELTYEN---YQSYSHRVSITTSFEIKSTGGIV--LGVPIFGWTHSSARNPFIY 509

Qy 526 SDKITQIPVVKAHTLVSGTTVIKPGP-FTGNNILKRTSSGPLAY---TSVSVKSPLSQRY 581

Db 510 ATKISQIPINKASRTSGAVWNFBGLYNGGPVKCLSGSGSQVINLRVATDAKG-ASORY 568

Qy 582 RARIYASTTNRLFWTIS-----GTIYISINVNKTMKGGDLLFTWTFDLATIGTA-F 633

Db 569 RIRIRYASDRAGKF--TISSSPENPATYSGASIAYTWMTNSTASLTYSTPAYAESGPINL 626

Qy 634 TFSNYSLSLTVGADSPASGCGEVVVDKPELIYPVNATPEABEDLDVAKKANGLFTSKDAL 693

Db 627 GISGSRTFDISITKEAANLYIDRIEFIPVNLFEABEDLDVAKKANGLFTNKDAL 686

Qy 694 QTSVTDYQVNAANLVECLSDLELYPNPKRMMLDVAKEAKELVQARNLLQDTGFNRINGEN 753

Db 687 QTSVTDYQVNAANLIECLSDLELYPNPKRMMLDVAKEAKELVQARNLLQDTGFNRINGEN 746

Qy 754 GWTGSTGIEVABGDVLFKDRSLRLTSAREIDTETPTLYQQIDESLLKPYTRYKLKGF 813

Db 747 GWTGSTGIEVVEGDVLFKDRSLRLTSAREIDTETPTLYQQIDESLLKPYTRYKLKGF 806

Qy 814 GSSODLEIKLIRHRANOYVKNVPDNLLPDVLPVNSCGIIDRCSEQQVVDANLALENNGEN 873

Db 807 GSSODLEIKLIRHRANOYVKNVPDNLLPDVRPVNSCGGVDRCSBQQVVDANLALENNGEN 866

Qy 874 GNMSDDSHAFSFHIDTGEIDLNENTGYIWVFKIPTNYGATLGNLELVERGPSLGSETLER 933

Db 867 GNMSDDSHAFSFHIDTGEIDLNENTGYIWVFKIPTNYGATLGNLELVERGPSLGSETLEW 926

Qy 934 AQOEOQWQDMARKRGASKAYAAKQAIDRLFPDYQDOKLNSGVEMSDMLAAQNLVQS 993

Db 927 AQOEOQWQDMARKRGASKAYAAKQAIDRLFPDYQDOKLNSGVEMSDMLAAQNLVQS 986

Qy 994 IPVYNDALPEIPGMNTYTFTELNRLOQAWNVLDLNAIPNGDFRNLGSDMNATSDDNV 1053

Db 987 IPVYNDALPEIPGMNTYTFTELNRLOQAWNVLDLNAIPNGDFRNLGSDMNATSDDNV 1046

Qy 1054 QQLSDTSVLVIPNMNSQVSQQFTVQPNYRYVLVARTAKEGVDGYVIIRDGANQTETLTP 1113

Db 1047 QQLSDTSVLVIPNMNSQVSQQFTVQPNYRYVLVARTAKEGVDGYVIIRDGANQTETLTP 1106

Qy 1114 NICDDDTGVLSADQTSYITKYTFPTSPTEQWIDMSETGVFNIESVELVLEE 1167

Db 1107 NICDDDTGVLSADQTSYITKYTFPTSPTEQWIDMSETGVFNIESVELVLEE 1160

RESULT 2
S49247
parasporal crystal protein cry9Cal [validated] - Bacillus thuringiensis
N;Alternate names: parasporal crystal protein cryIH
C;Species: Bacillus thuringiensis
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 01-Dec-2000
C;Accession: A59350; S49247
R; Lambert, B.; Buysse, L.; Decock, C.; Janssens, S.; Piens, C.; Saey, B.; Seurinck, J.; Vervaeke, W.
Appl. Environ. Microbiol. 62, 80-86, 1996
A;Title: A Bacillus thuringiensis insecticidal crystal protein with a high activity against corn borers
A;Reference number: A59350; MUID:96141404; PMID:8572715
A;Accession: A59350
A;Molecule type: DNA
A;Residues: 1-1157 <L>
A;Cross-references: EMBL:Z37527; NID:G547554; PIDN:CAA85764.1; PID:G547556
A;Experimental source: serovar tolworthi
C;Comment: This parasporal crystal protein, active against corn borer and other insects,
C;Superfamily: parasporal crystal protein
C;Keywords: delta-endotoxin

Query Match 51.7%; Score 3123; DB 1; Length 1157;

Db 995 IPGIVEIYTELSDRLQQAQSYLTSRMAVQKGFNSGLDSWNTMTDASVQDQGNMHPVL 1054

Qy 1065 PNNNSVQSQFTQPNRYRYLVRTARKEGVGDGVIIIRDGANOTETLTFNICDDDTGVL 1124

Db 1055 SHWDAQVQQLRVNPNCKYLVRTARVGGGQGVYIIRDCGHHQETLTFNACDYNVTY 1114

Qy 1125 ADQTSYITKTVETFPSTPEQWIDMSETEGPNFNRISVELVLEE 1167

Db 1115 VNDNSYITEVVPYETKMMVSESEGSFYIDSEFIETOE 1157

RESULT 3

S39536

parasporal crystal protein cry9Ba1 - Bacillus thuringiensis

N;Alternate names: delta-endotoxin-related protein; parasporal crystal protein cryX

C;Species: Bacillus thuringiensis

C;Date: 07-Oct-1994 #sequence_revision 01-Dec-1995 #text_change 01-Dec-2000

C;Accession: S39536

R;Shevelev, A.B.; Svarinsky, M.A.; Karasin, A.I.; Kogan, Y.N.; Chestukhina, G.G.; Stepan

FEBS Lett. 336, 79-82, 1993

A;Title: Primary structure of the cryX, the novel delta-endotoxin-related gene from Baci

A;Reference number: S39536; MUID:94085596; PMID:8262221

A;Accession: S39536

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1154 <SHE>

A;Cross-references: EMBL:X75019

C;Superfamily: parasporal crystal protein

C;Keywords: delta-endotoxin

Query Match 45.5%; Score 2750.5; DB 2; Length 1154;

Best Local Similarity 48.7%; Pred. No. 1.9e-151;

Matches 562; Conservative 197; Mismatches 341; Indels 53; Gaps 18;

Qy 41 NYKDYLRMSGEENDELFCNPEFTFSSSTVQTCIGVQVIGALGVPPAGQIASFVSFVG 100

Db 29 SYKDYLRMSGEYDYSINPGNV--RTGLQGTIDIVAVVVGALGPGVGGILTGLSTLFG 86

Qy 101 QLWPSSTVSVNEMIMKQVEDLIDOKITDSVRKTLALAGLQGLDGLVYQKSLKWLNRN 160

Db 87 FLWFSNDQANWEAFIEQHEELIEQISDVVRLTALDGLTQNYNYQLIAKWEERPN 146

Qy 161 DTRARSVVVQYIALELDLFAKIPSPALSGQ-----EVLPSVYQAQANLHLLLRASI 215

Db 147 GVRA-NEVLQRFETLHALFVSSMPSFG-SGPGSQRFQALLVYVYQAQANLHLLLRDAEK 204

Qy 216 FGAENGTPGEISFYDROV--TRTAQYSDYCKVNTVGLDKLKTNAASWLKYHOFREEM 274

Db 205 YGARWGLRESQIGNLYFNEQLQTRDRTDYNHCVNAYNNGLAGLRTSAESWLKYHOFREA 264

Qy 275 TLLVLDLVALPPNYDTRTYPIETTAQLTREYVTDPIVFNRETSGF----CRRWLSNDSI 330

Db 265 TLMAMDILALEPYNTNRYPIAVNPQLTREYVTDPLGVPSSESLFPELRLCLRINQWETSAM 324

Qy 331 SPFSVESAVIRASPHLFDILSIBRIFYTRAGLPLNNTLEYLVVGHSI--KYKNTNASSAL 388

Db 325 TFSNLENAIISSPHLEFDITINNLMYTCFSVHLTN-QLIEGWIGHSVTSLSLGSPTVL 383

Qy 389 ERNYGTTISNKIKYVDLANKDIFQVRS---IGADLANVYAOVYGVYPYASFTLLDRNTGSG 445

Db 384 RRYNGSTTS-IVNYSFSDRDVYQINTSRHTGLGFQN--APLPGITRAQF----- 430

Qy 446 SVGGFTYSKPHTTMQVCTQNTYNTIDEIP--PENEPLSRGYSHRLSHITSYSF-----SK 497

Db 431 -YPGTTYSVTQRNALTCQNTNSIDELSPDNPEFSKYSYSHRLSHITSYLRVLTIDGI 489

Qy 498 NASSPARYGNLPVFAWTHRSADVNTVYSDKIQIPIPVVKAHLTVSGITVIKPGPFTGNI 557

Db 490 NIYS-----GNLPTVWTHRDVLTNTIADRTQLPLVKSFEIPAGITVVRGPGFTGDI 545

Qy 558 LKPTSSGPLATYSVSKPQLSQRYARIRYASTTNLRLFTVITSGTRIYSINNTKMKGD 617

Db 546 LRRTGVTGFTIRVIRTAFLQRYIRERFPASITNLFLGIRVGRQVNYFQFGRTMNRGD 605

Qy 618 DLTFNTFDLATICTAFTFSNYSDSLTVGADSPFASGGRYVVDKFKELIPVNTATFEAEEDLV 677

Db 606 ELRYESPATREFTTDFNFRQQLISVFANAFSAGQEVYFDRIBIIPVNPAREAKEDLEA 665

Qy 678 AKAVNGLPFSKDALQTSYTDYVQAANLVECLSDLELYENKRMMLWDVAKRKLVA 737

Db 666 AKKAVASLFRTRDGLQVNVKQYDQAANLVCLSDLEQYGYDKMLLEAVRAAKRLRE 725

Qy 738 RNLQDTGFNRING--ENGWGTGTGIEVAGDVLFDORSLSLTSAREIDTSTPTLYLQ 795

Db 726 RNLQDQDFNTINSTENGWKAANGVTISEGGPYKGRALQASAR-----ENYPTIYQK 781

Qy 796 IDESLKLPYTRYKLGFIGSSQDLLEIKLIRHRANQIVNVPDNLPLDVLVNSCGGIDRC 855

Db 782 VDASELKPYTRYRSRSDGFPVKSQDLLEIDLIEHHKVLVKNVPDNLVSDTYPDDSCGINRC 841

Qy 856 SEQDYDANLALNNGR-NGNMSDSHAPSHIDTGEIDLNENTGIWVVKPIPTNGYAT 914

Db 842 QBOQMNAQLTEHHHPMDCCRAAQTHFESYIDTGLNSSVDQGIWAIFKVRTIDGVT 901

Qy 915 LGNLELVEEGPLSGETTLERAAQQOQWQDKMARKGASEKAYYAAQKQIDRLFADYQDQK 974

Db 902 LGNLELVEGVLGSGESLEREQDNTKSAELGRKAETDRVYQDAKOSINHLFVDYQDQ 961

Qy 975 LNSGVMSDLAAQNLVOSIPYVYNDALPEIPGANYTSFTBLTRLOQMMLYDLRNAIP 1034

Db 962 LNPEIGMADIMDAQNLVASISDVYSDAVLQIPGINYEIYTELSNRLQAQSYLYTSRNAVQ 1021

Qy 1035 NGDFRAGLSDNATSDVNVQQLSDTSVLVPPNNSQVSOQFTVQPNRYVYVLRVARTKGV 1094

Db 1022 NGDFNGLSDGNATAGASVQDQGNTHFLVLSHWDQAQVSQQPRVQPNCKYVLRVARTKGV 1081

Qy 1095 GDGVYIIRDGANOTETLTFNICDDDTGVLSDAQTSYITKTVETPSTBQWIDMSBETGV 1154

Db 1082 GDGVYIIRDGANHTETLTFNACDIDINGTYVDNTYLTKEVIFYSHTEDMVEVNETGA 1141

Qy 1155 FNIESVELVLEE 1167

Db 1142 FHIDSIEFVETEK 1154

RESULT 4

S00873

parasporal crystal protein cryBa1 - Bacillus thuringiensis subsp. thuringiensis

N;Alternate names: parasporal crystal protein cryA4

C;Species: Bacillus thuringiensis subsp. thuringiensis

C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 01-Dec-2000

C;Accession: S00873

R;Brizzard, B.L.; Whiteley, H.R.

Nucleic Acids Res. 16, 2723-2724, 1988

A;Title: Nucleotide sequence of an additional crystal protein gene cloned from Bacillus

A;Reference number: S00873; MUID:88203216; PMID:3362680

A;Accession: S00873

A;Molecule type: DNA

A;Residues: 1-1228 <BRI>

A;Cross-references: EMBL:X06711; NID:940264; PIDN:CAA29898.1; PID:9580949

C;Genetics:

A;Gene: cryA4

A;Start codon: TTG

C;Superfamily: parasporal crystal protein

C;Keywords: delta-endotoxin

Query Match 42.1%; Score 2547.5; DB 2; Length 1228;

Best Local Similarity 43.8%; Pred. No. 1.3e-139;

Matches 558; Conservative 183; Mismatches 381; Indels 151; Gaps 21;

Qy 1 MSPNNQYEYELDASSSTSVSDNSVRYPLANDQTTLQNMNYKYDLRMSSEGNPELFGNP 60

Db 1 MTSNRKNEIINAVSNHSA-----QMDLLPDARIEDLSICIAEG-----NNI 42

Qy 61 ETPISSTVQTCIGVIGVQVIGALGVPPAGQIASFVSFVIGQVLPSTSVSVNEMIMKQVED 120

Db 43 DPFVASTVQTGINIAGRIILGVLPFAGQASFYSLVGLWLPGR-DOMEIFLHEVQ 101
Qy 121 LIDOKITDSVRTALAGLQGLDGLDYYQSKLKNWLENDRTRASVVVVTVQVIALELDFV 180
Db 102 LINQITENARTALARLQGLGDSFRAQQSLEDWLENRDDARTSVLYTVQVIALELDFL 161
Qy 181 AKIPSPALSGQEVPLLSVYQAANLHLLLDASIFGAEWFTPGHISTFTFDQVTRTAQ 240
Db 162 NAMPLFAIRNQEVPLLVYQAANLHLLLDASIFGSEFGLTSQBIQRYERQVTRD 221
Qy 241 YSDYCVKMYNTGLDKLKTNAASMLKHOPRREMTLLVLDLVALPVPNDTRTPYLETAAQ 300
Db 222 YSDYCVKMYNTGLSLRGTNAASWRYNQFRDLGLVLDLVALPVPNDTRTPYLETAAQ 281
Qy 301 LTREYVTDPIVFNRETSGGFCRWSLNSDIISFSEVESAVIRSPHLDILSBIFFYTRAG 360
Db 282 LTREYVTDPIVFNRETSGGFCRWSLNSDIISFSEVESAVIRSPHLDILSBIFFYTRAG 337
Qy 361 LPLNTEVLEWVGHISIKYK-----NTNASSALERNYGTITSNKIKYDILANKDIFQVR 414
Db 338 SRWSNTRHMTYWRGHTIQSRPIGGGLNTSTHGATNTSINPVLK-----FASRDVYRTE 391
Qy 415 SLGADL--ANTYAAOVYGVYPAFTLLDKNTGSGVGGFTYKSPHTTMQVCTONVNTIDBI 472
Db 392 SYAGVLLWGIYLEPIHGVPTFRFNTFNQNTSDR--GTANYQPYESPGLQKDSF--EL 448
Qy 473 PPE--NEPLSRGYSHRLSHITSYSPSKONASPARYGNLPVPAWTHRSADVNTVYSDKIT 530
Db 449 PPEITERPNYSYSHLSHIGILQSR-----VNVVYVSWTHRSADRTNIGPNRIT 500
Qy 531 QIPVKAHTLVSGTIVKGPFTGQGNILKRTSSGPLATYSVSPLSQRTARIRYAST 590
Db 501 QIPVKAHTLVSGTIVKGPFTGQGNILKRTSSGPLATYSVSPLSQRTARIRYAST 560
Qy 591 TNLRLPVTISCTRIYSINVNKTKNGDDLTFNTPLDITAGTAPTSYSDSLTVGADSA 650
Db 561 VDFPFVSRGGTIVNPFRLNTSGDELKYNFVRRAFTPTPTTQIQDILRISIQGLS 620
Qy 651 SGGEVYVDKFLIPVNAATFEEDLDVAKQAVNGLFTSKD--ALQTSVTDYQVNAANLV 709
Db 621 NGGEVYVDKFLIPVNAATFEEDLDVAKQAVNGLFTSKD--ALQTSVTDYQVNAANLV 680
Qy 710 ECLSDLYPNKRMKMDVAKRLVQARNLLQDTGPNRIN----- 750
Db 681 ACLSDEFCLDEKRELLEKVKAKRLSDERNLLQDPNFTSINKQDPFISTNBSQNFSTS 740
Qy 751 -GEGWNTSGTGEVAGDGLVLFKDRSLTSAREIDTETPTLYQOIDESELLKPYRYKL 809
Db 741 QSEKGNWSENITQEGNDVFAKENVTLPGT---FNECYPTTLVQKIGESLKAATRYQL 797
Qy 810 KGFIGSSQDLKILIRHRANOIVKNVP--DNLLPDVL--PVNSCGIDRCSEQQVVDANL 865
Db 798 RGYIEDSQDLKILIRYNAKHEBTLDPGCTESLWPLSVESPIGRCEPNRCA--PHFEWNP 855
Qy 866 ALNENGNG--NWSDSHAFSPHIDGEBDLENCTGIWVFKIPTNGYATLGNLELVEEG 924
Db 856 DLDCSRDGEKCAHSHFSLDIDVGCCTDLHNLGVWVVFVKIQBEGHARLGNLEFIBEK 915
Qy 925 PLSGTLLERAOQO 984
Db 916 PLLGALSRLVKAERKAKKREKLEKLEKLEKLEKLEKLEKLEKLEKLEKLEKLEKLEKLE 975
Qy 985 LAAQNLVOSIPVYNDALPEIPGMNYTSTFELTNLQOANWLYDLRNALPNGDFPAGLSD 1044
Db 976 HAADKLVRHIREAYLSLFPVPGVNAEIPFEELEGHIIITAIISLYDARNVVKNGDFPAGLTC 1035
Qy 1045 WNAISDVNVQQLSDTSLVLIPIWNSQVQOQPTVQPNRYVLRVTRARKEGVDGYIIRDG 1104
Db 1036 WNVKGVDPVQOQSHRSHSLVPIPEWAEVSAVRVCPGCGYILRVATYKRGYEGGCVTIHEI 1095
Qy 1105 ANOTETLTP----- 1113
Db 1096 ENNTDELKFKRHEEVPTDTGTCTNDYTAHQGTAGCADACNSRNAGYEDAYEVDTTASV 1155

Qy 1114 -----NICDDTGVLSAD--QTSYITKTVEFTSTBQVWIDMSETRGV 1154
Db 1156 NYKPYEBETVTDVRDNHCEYRGVNVYPPVAGVYTKELFYFPETDTVWIEIGTEGK 1215
Qy 1155 FNIESVELVLEER 1167
Db 1216 FIVDSVELLMEE 1228
RESULT 5
A48944
paraesporal crystal protein cry7Aa1 - Bacillus thuringiensis
N:Alternate names: paraesporal crystal protein cryIIIC
C:Species: Bacillus thuringiensis
C>Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 01-Dec-2000
C:Accession: A48944
R:Lambert, B.; Hofte, H.; Anny, K.; Jansens, S.; Soetaert, P.; Peferoen, M.
Appl. Environ. Microbiol. 58, 2536-2542, 1992
A>Title: Novel Bacillus thuringiensis insecticidal crystal protein with a silent activi
A:Reference number: A48944; MUID:92384571; PMID:1514800
A:Contents: B7S1373
A:Accession: A48944
A>Status: preliminary
A:Molecule type: DNA; protein
A:Residues: 1-1138 <LAM>
A:Cross-references: GB:M64478; NID:g142760; PIDN:AAA22351.1; PID:g142761
A>Note: sequence extracted from NCBI backbone (NCBIN:112092, NCBI:112093)
C:Superfamily: paraesporal crystal protein
C:Keywords: delta-endotoxin

Query Match 40.7%; Score 2458.5; DB 2; Length 1138;
Best Local Similarity 43.5%; Pred. No. 1.6e-134;
Matches 514; Conservative 208; Mismatches 400; Indels 59; Gaps 17;

Qy 1 MSPNNQNEYILDASSSTVSNSVRYPLANDQTTTLQNMNYKDYLRMSGEBNPELFGNP 60
Db 1 MNLNLDGTE-----DSNRTLANSNLYTQKALSPLKMNQYQDLSITRERQPEALASG 55
Qy 61 ETFISSSTVQTGIGIVGQVLAGVPFAGQIASFYSLVGLWLPSTVSVMKMKQVED 120
Db 56 NT-----AINTVSVTVGATLSALGVGASFTNPKYIAGLLWPENG-KIMDFNTEVEA 109
Qy 121 LIDOKITDSVRTALAGLQGLDGLDYYQSKLKNWLENDRTRASVVVVTVQVIALELDFV 180
Db 110 LIDOKIEYVNAKAEIADGLGSLADRYKQALADWLGKDDPEALISVATEPRIIDSLPE 169
Qy 181 AKIPSPALSGQEVPLLSVYQAANLHLLLDASIFGAEWFTPGHISTFTFDQVTRTAQ 240
Db 170 FSMSPKVTGYEIPLLTVYQAANLHLLLDASITLYGDKNGFTQNNIBENTYRQKRIS 229
Qy 241 YSDYCVKMYNTGLDKLKTNAASMLKHOPRREMTLLVLDLVALPVPNDTRTPYLETAAQ 300
Db 230 YSDHCTKWNYSGLSRLANGTYEQMNTYNNRFRREMLMALDLVAVPPFDPFRYSMETSTQ 289
Qy 301 LTREYVTDPIVFNRETSGGFCRWSLNSDI--SSEVESAVIRSPHLDILSBIFFYTR 358
Db 290 LTREYVTDPIVFNRETSGGFCRWSLNSDI--SSEVESAVIRSPHLDILSBIFFYTR 338
Qy 359 AGLPLNTE-YLEYVWGHISIKYKNTNASSALERNYGTITSNKIK--YVDLANKDIFQVR 415
Db 339 YKASHEIQDPLFWSAHKVSFKSEQSNLYTGYGTSGYISSGAVSPHGNIDIVRTLA 398
Qy 416 LGADLANYYAQVGVYPAFTLLDKNTGSGVGGFTYKSPHTTMQVCTONVNTIDBIPE 475
Db 399 APSVWVYPTQNYGVEQVEFY-----GVKGVHVRGDKNYDL---YDSDIDQLPPD 446
Qy 476 NEPLSRGYSHRLSHITSYSPSKONASPARYGN--LPVPAWTHRSADVNTVYSDKITQIP 533
Db 447 GEPHIEKYTHRLCHATAI-----PKSTPDYDNATIPFISWTHRSABEYNNRIPNKLITIP 501
Qy 534 VVKAHTLVSGTIVKGPFTGQGNILKRTSSGPLATYSVSPLSQRTARIRYASTTNL 593

Db 502 AVKMYKLDPPSTVVKPGFTGCDLVKRGSTGYIGDIKATVNSPLSQYRVVRVATNVSG 561
Qy 594 RLFTYISG-----TRIYSINNVKTMKGGDLTFNFDLATIGTAFTSNTSYSDSLTVGADS 648
Db 562 QFNYYINDKITLQTKFC--NTVETIGSGKDLTSGFGYBYSTTIOFPDEHPKITLHSD 619
Qy 649 FASGGEVYVDFELIPVNAFPEABEDLVAKAVNGLPTSKDALQTSVTDYQVNOAANL 708
Db 620 LSNSSFFVDSIEPIPDVNVYAEKEKLEKAKAVNTLFTGKNALQNDVDYKQVDSIL 679
Qy 709 VECLSDLEYPNKRLMDVAKRELVAARNLLQDTGFNRING--ENGWGTGSGTIEVAEG 766
Db 680 VDCISGDLYPNKKELQNLVYAKRLYSRNLILDPTFDSINSEENGWYSGNGLVING 739
Qy 767 DVLFKORSKLTSAEIDTETPTLYLQQIDESLLKPTTRYKLGKPTGSSODLEIKLIRH 826
Db 740 DFVFKGNLYLIFSGTN--DTQ-YPTLYQKIDESKLEKKEYTRYKLGKPTGESSQDLAEVYIRY 796
Qy 827 RANQIVKGVDPNLLPDLVFNSSCGIDRCSEQQVVDANLALENNGENGMSSDSHAFSPH 886
Db 797 DAKHRTLVDSNLLPDLIPENTCEPNKCAQQVLDENPFCSSMQDGLSDSHSPSLN 856
Qy 887 IDTGEIDNENGTGWVFKIPITNGYATLGNLELVBERGLSGTETLERAQQEQEQMDKMA 946
Db 857 IDTGSINHNENLGIWLFKISTLEGYAKFKNLEVEIDGPIVGEALARKVQETKRNKLA 916
Qy 947 RKGAASEKAYAAQALIDRLPADYQDQKLSGVEMSDMLAAQNLVQSPVYNDALPEIP 1006
Db 917 QLATTETQALYTRAKQALDNLFAQAQSHLRDVTFAEIAARKIVQSIREAYMSLWVWP 976
Qy 1007 GMYVTSFTELNLQRAWNLVDLENALPNGDFRGLSDNATSDVNVQQLSDTSVLVPIPN 1066
Db 977 GVNHPIFELSGVRQAFQYDVNVNRGRFLNGLSDWIVTSVDRVQBERGNVNLVNN 1036
Qy 1067 WNSQVSQOFTVQPNRYVLRVATRKEGVGDGYVIRDEGANQTETLTFNICDDDTGVLSAD 1126
Db 1037 WDAQVLQNLVLYQDRGVLVHTARKIGIGSGYITIDEEGHTDQLRFTAC-REIDASNAF 1095
Qy 1127 QTSYITKVTPEFTSTEQVMDMSETEGVNFVNSVELVLEE 1167
Db 1096 ISGYITKLELFFPDTEKVVHIEIGTEGIFLVSIELFLMBE 1136

RESULT 6
S19306
paraaporal crystal protein cry9Aa1 - Bacillus thuringiensis
N;Alternate names: Delta-endotoxin; insecticidal crystal protein; paraaporal crystal pro
C;Species: Bacillus thuringiensis
C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 01-Dec-2000
C;Accession: S19306; S23588; A44847; S14602; S14837
R;Smulevitch, S.V.; Osterman, A.L.; Shevelev, A.B.; Kaluger, S.V.; Karasin, A.I.; Kadyrc
FBS Lett. 293, 25-28, 1991
A;Title: Nucleotide sequence of a novel delta-endotoxin gene cryIG of Bacillus thuringie
A;Reference number: S19306; MUID:92070568; PMID:1660003
A;Accession: S19306
A;Molecule type: DNA
A;Residues: 1-1156 <SMU>
A;Cross-references: EMBL:X58120; NID:9870929; PIDN:CAAA41122.1; PID:g40271
A;Experimental source: subsp. galleriae
A;Accession: S23588
A;Molecule type: protein
A;Residues: 24-34 <SMU>
A;Experimental source: subsp. galleriae
R;Gleave, A.P.; Hedges, R.J.; Broadwell, A.H.
J. Gen. Microbiol. 138, 55-62, 1992
A;Title: Identification of an insecticidal crystal protein from Bacillus thuringiensis
A;Reference number: A44847; MUID:92211329; PMID:1556556
A;Accession: A44847
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1151 <GLE>
A;Cross-references: EMBL:X58534; NID:g48879; PIDN:CAAA41425.1; PID:g48880
A;Experimental source: isolate DSR517

A;Note: sequence extracted from NCBI backbone (NCBIN:92865, NCBIP:92867)
C;Genetics:
A;Gene: cryIG
C;Superfamily: parasporal crystal protein
C;Keywords: delta-endotoxin

Query Match 36.3%; Score 2191; DB 2; Length 1156;
Best Local Similarity 41.5%; Pred. No. 5.7e-119; Indels 98; Gaps 24;
Matches 501; Conservative 195; Mismatches 414;

Qy 5 NQNEYELDASSVSVDNSVRYPLANDQTTTLQNMNYKQYLRMSEGENPELFGNPETPI 64
Db 2 NQNKGGIIGASGCCASDDVAKYPLANNPYSSALNIN-----SC 40

Qy 65 SSSTVQTGIGIVGQ-----VLGALGVP-FAGQIASFYFTVQQLAPSSVSVWE 112
Db 41 QNSILANWINIIGDAEKAESVIGTTIVSLITAPSLTGLISIVYDLIGKVLGSGSQSISD 100

Qy 113 MIMKQVEDLIDQKITDSVRKTAALAGLQGLDGVYQKSLKNWLENRNDTRARSVVVYQY 172
Db 101 LSIQDLUSIIDLRVSQSLVNDGLADFNQSVLLYRNYLEALDSWKNKNPNSASABELRTRPR 160

Qy 173 IA-LELDFVAKIPSPFAISGQ-----EVPLLSVYAQAANLHLLLRDASIFGAEMGP --- 222
Db 161 IADSEFDRILTRGSLTNGSSILARQNAQIILLPSASAAPFHLILLRLDRTATRYGTNGLYNA 220

Qy 223 TPGEISTPYDROVTRTAQYSDYCVKWTNTGLDKL--KGTNAASWLKHYQPREMTLLVLD 280
Db 221 TP--PINYQSKVELIBLYTDYCVHMYNRGPNELRQRTSATAWLEPHRYREMTLMVLD 278

Qy 281 LVALPNYDTRTPIETTAQLTREVVYDPIVFNRETSGPCRRWSLNSD-----ISP 332
Db 279 IVASFSSLDITNTPIETDFQLSRVYITDPI-----GPHRSLSRGESWFSFVNRAFP 330

Qy 333 SYVESAVIRSEPHLFDILSIEFYTTTRAGLPLN-NTEYLEYVWGHISIKYKTNASSALERN 391
Db 331 SDLENA-IPNPRSPFWLNNMIISTGSLTLPVSPSTDRAVWYGSRRDISPANSQPIITBLI 389

Qy 392 YGITSNKIKYDOLANKDIFQVRSGLADLANYYAQVGVYPYASFTLDDKRTGS-----G 445
Db 390 SQQHTT---ATQITLGERNIFRVDQSACNLND---TTYGVNRAVP-YHDASGSGRSRYVEG 442

Qy 446 SVGGFTYSKPHTTMQVCTQNYNTIDEIPPENE--PLSRGYSHRLSHITSYFSKNASPA 503
Db 443 YIRTTGIDNPR-----VQNTY--LPGENSDIPTEDYTHILSTINLTGGLRQVASN 494

Qy 504 RYGNLFPVFAWTHRSADVNTVYSDKITQIPVKAHTLVSGTTVTKGPFCTGNNILKRTSS 563
Db 495 RRSSLVWYGTWTHKSLARNNTINPDRIITQIPLTKVDTRTGTGVSYYVNDPGFICGALLQRTDH 554

Qy 564 GPLATYSVSVKSPLSQRYRARIYASTNRLPFTIIS-GRYIYSINNVKNTNWKGGDLTFN 622
Db 555 GSLGVLRVQPPHLRQOQYRIRVYASTNIRLSVNGSFGT--ISQNLPTNRLGSDLYAG 612

Qy 623 TFDLATIGTAFPSNYSDSLTVCADSPASGGEVVDKPFELIPVNAFPEABEDLVAKKAV 682
Db 613 SPAIREPNTSIRPTASPDQRLTFTBPSFIRQEVTVDRIBFIPVNPTRBEKDLERAKAV 672

Qy 683 NGLFTSKDALQTSVTDYQVNOAANLVECLSDLEYPNKRLMDVAKRELVAQARNLLQ 742
Db 673 ASLFTTRDGLQVNVKDYQVDQAANLVSCLSDEQYGVYDKMLLEAVRAAKLSRERNLLQ 732

Qy 743 DTGFNRING--ENGWGTGSGTIEVAEGDLVKORSKLTSAEIDTETPTLYLQQIDESL 800
Db 733 DPDEFNTINSTEENGWAKASGVITSEGGPFYKRAIQLASAR-----ENYPTIYQKVDASE 788

Qy 801 LKPYTRYKLGKPTGSSQDLLEIKLIRHANQIVKNNVDPNLLPDLVFNSSCGIDRCSEQY 860
Db 789 LKPYTRYRDLGFPVKSSQDLLEIKLIRHANQIVKNNVDPNLLPDLVFNSSCGIDRCSEQY 848

Qy 861 VDNALALENNGE-NGNMSDSDSHAFSPHIDTGEIDNENGTGWVFKIPITNGYATLGNLE 919
Db 849 VNAQLETEHEHPMDCCCEAAQTHFSSVIDTGDNLSSVDQGIWAIFKVRTTDGTGATLGNLE 908

Query Match 34.6%; Score 2093.5; DB 2; Length 1166;
Best Local Similarity 39.7%; Pred. No. 2.7e-113;
Matches 482; Conservative 172; Mismatches 392; Indels 167; Gaps 24;

QY 59 NPETFISSVTQIGIVGOVLGAL-----GYPPAGQIASFYFVGOGLWPSSTVSVM 111
DB 17 NPSEIFNAR-NSNFGVLVSQSSGLTRPFLLEAAVPEAGFALGLFDII-----WGLGVQDW 71

QY 112 EMIMKQVEDLDQKITDSVRTALAGLDGDLVYQKSLKWLNRNDRASVVVVTQ 171
DB 72 SLFLRQIEQLIRQITELERNRATAILTGLSSSYNLVYEAALREWENDPNNPASQVRTR 131

QY 172 YIALELDPVAKIPSAISQGEVPLLSVYQAANLHLLLRDASIFGAENGFTPGSEISTFY 231
DB 132 FRUTDDAIVTGLPFLAIRNLEWNLVSVYTOAANLHLLSLRDVAVYFGRWGLTQANIEDLY 191

QY 232 DRQVTRTAQYSDYCVKWNVTGLDKLKGNTAASWLYKHQFRREMTLLVLDLVALFPNYDTR 291
DB 192 TRLTNSIOEYSDHCARWYNQGLNEIGISR-----RYLDFQFDLTISVLDIVALFPNYDIR 247

QY 292 TYPIETTAQATREYVTDPIVFNRETSGGFCRRWSLNSDIFSSEVESAVIRSPHFLDILSE 351
DB 248 TYPIETQSLTREITYTSPV-----AGNINFGLSIANVLRAPHLMDFIDR 292

QY 352 IEPTTTRAGLPLNTEYLEYVWGHISIKYKNTNASSALERN---YGT----- 394
DB 293 IVIYT-----NSVSTPYWAGHEVISRRTGQGGNEIRFPPLYGVAANAABPPVITIRPTG 345

QY 395 ITSNIKIVYDLANKDIFQVRSIGADLANVYAQVYGVYASFTLMDKNTGSGVGFT-YS 453
DB 346 FTDEQRQWY-RASRVVSPRSSGD-----FSLVD-----AVGFLTIPS 383

QY 454 KPHTTMQVCTONVNTIDEIPPENEPLSRGYSRLSHITSYSPSKNASSP--ARYGNLPVP 511
DB 384 AVSIYRNGFGFNTDTIDEIPIEGTDPPTGYSHRLCHVGFL-----ASSPPTSOVARAPI 438

QY 512 ANTHRSADVNTVYSKIQIOPVVKAAHTLVSGTIVKGPFTQGNILKRTSSGPLATSV 571
DB 439 SWTHRSATLNTIAPDVIQIPLKAFNLHSGATIVKGPFTGCDILRRTNVGSGDMRV 498

QY 572 SVKSPLSORVRRIRYASTTNLRFLVTISGTRIYSINVNKMNGDDLTNTFDTLATIGT 631
DB 499 NITAPLSQRVRIRYASTTDLQPTTNINGTTINGNFSTMGDDLOQYGRFVAGFTT 558

QY 632 APTFSNYSDSLTVGADSFASGGEYVDKFLIPVNATFEABEDLDVAKAVNGFLTSGD 691
DB 559 PPTFSDANSTFTIGAFGSPNNEVYIDRIEFPVPAEVTFEARYDLEKAKAVNALFTSSNQ 618

QY 692 -ALQTSVTDYOVQOANLVECLSDLYPNKRMKMDAVKAKRLVQARNLLQDTGNRIN 750
DB 619 IGLKTDVTDYHDKVSNLVECLSDLYPNKRMKMDAVKAKRLVQARNLLQDTGNRIN 678

QY 751 --GENGWTSGLIEVAGSDVLEKDRSLRLTSAREIDETPTLYLQOIDESESLKPYTRYK 808
DB 679 RQDRGRGSDTITQGGDVFKEVNYVLPCTFD---GCYPTLYLQKIDSKLVYTRYQ 735

QY 809 LKGPISGSQDLEIKLIRHRANQIVKVPDN-----LLPDVLPVNSCGGIDRCSEQQYVDAN 864
DB 736 LRGVIEDSQDLEIYLIRYNAKHETVNVPGTGLWPLSAQSPIGKCGEPNCA--PHLEWN 793

QY 865 LALENGENG-NMSSDSHAFSFHIDTGEIDLNENGTGWVFKIPTTNGYATLGNLELVRE 923
DB 794 PDLDCSRNGEKAHSHHESLDIDVGCTDLNEDLGVWVIFKIQDGHARLGNLEPLEE 853

QY 924 GPLSGETLERAOQOQOQWEMARKGKASEKAYAAKQADRLPADYQDQKLSGVEMSD 983
DB 854 KPLLGEALARYKRAEKWRDKREKLELETINIVYKEAKESVDALFVNSQYDOLQADNTIAM 913

QY 984 MLAANLVQSIPIVYNDALPEIGPMYNTSFTELNRLOQAWNLVDLNAIPNGDFRNLGS 1043
DB 914 IHAADKKVHSIRREYLPVPGVNAAI FEELEGRIFTAFSLYDARNVINGDFRNLGS 973

QY 1044 DMNATSDVNV-QQLSDTSVLVIPNNNSOVSOQFTVQPNRYVLRVYTABKEGVDGWIIR 1102
DB 974 CMNVKGVHVEEQNNHRSVLVPEWBAEVSQVPCGCGHILRVTAYKEGVSQCVIH 1033

QY 1103 DGANQOTLTLPN-----ICDDDTG-----VLSA 1125
DB 1034 EIENTNDELAFSNCVBEVYPNNTVTQNDYTANQEEYKGAVTSHNRGYDEAYGNPNPSPA 1093

QY 1126 DOT-----SYTKTVETPSTQWIDMSSETGV 1154
DB 1094 DYTPEYKEAYTDGRRENPCBSNRGYDTPPLPAGYVYTKLEYEYPPETDKVWIEIGETGT 1153

QY 1155 FNIESVELVLEE 1167
DB 1154 FIVESVELLME 1166

RESULT 9
A26513
parasporal crystal protein - Bacillus thuringiensis (strain aizawai)
C:Species: Bacillus thuringiensis
C:Date: 11-Mar-1988 #sequence_revision 11-Mar-1988 #text_change 01-Dec-2000
C:Accession: A26513
R:Oeda, K.; Ohie, K.; Shimizu, M.; Nakamura, K.; Yamamoto, H.; Nakayama, I.; Ohkawa, H.
Gene 53, 113-119, 1987
A:Title: Nucleotide sequence of the insecticidal protein gene of Bacillus thuringiensis
A:Reference number: A26513; MUID:87248103; PMID:3297927
A:Accession: A26513
A:Molecule type: DNA
A:Residues: 1-1155 <OE>
A:Cross-references: GB:M16463; NID:g143098; PIDN:AAA22551.1; PID:g143099
C:Superfamily: parasporal crystal protein
C:Keywords: Delta-endotoxin

Query Match 34.2%; Score 2064.5; DB 2; Length 1155;
Best Local Similarity 39.0%; Pred. No. 1.3e-111;
Matches 476; Conservative 184; Mismatches 387; Indels 173; Gaps 27;

QY 53 NPDL-----FENPET-FISSVTQGT---IGI---VGQVLGALGVPPAGQIASFYSP 97
DB 4 NPINECIPYCNLSNPEVLEGGRIETGVTPIDLSLSTQFLSEFPVGGAGFVLGLVDI 63

QY 98 IVGQLWPSSTVSWEIMKQVEDLDIQITDSVRKTALAGLDGDLVYQKSLKWL 157
DB 64 IWGIFGFSQ---WDAFLVQIEQLINQRIEFAFNQAIISRLGSLNLYQIVAESFWEA 119

QY 158 NRNDTRARSVVVTOYIALELDFVAKIPSAISQGEVPLLSVYQAANLHLLLRDASIFG 217
DB 120 DPTPALREKRIQFNDMSALITAIPLFAVQNVQVPLLSVYQAANLHLSVLRDVSVFG 179

QY 218 AEMGFTPEISTFYDRQVTRTAQYSDYCVKWNVTGLDKLKGNTAASWLYKHQFRREMTLL 277
DB 180 QRWGFDAATINSRYNDLTILIGNYTDHAWRYNTGLERWVGDPDSRDWIRYQFRELTLT 239

QY 278 VLDLVALFPNYDTRTYDIETTAQATREYVTDPIVFNRETSGGFCRRWSLNSDIFSSEVES 337
DB 240 VLDIVSLFPNYDSTRYDTRTYDIETTAQATREYVTDPIVFNRETSGGFCRRWSLNSDIFSSEVES 285

QY 338 AV---IRSPHLFDLSIEFVYTRAGLPLNTEYLEYVWGHISIK---YKNTNASSALE 389
DB 286 GIGSGIRSPHMLDILNSITVT-----DAHGEY---YWSGHQIMASPVGSPGPTFPL- 337

QY 390 RNYGTTISNKIKYVDLANHDIPOVRSIGADLANVYAQVY--GVYASFTLLDKNT----G 443
DB 338 --YGTMGNAAPQQRIVAIQCGVYVTRLSSTL---YRRPFNIGINNQQLSVLDGTGFAYGT 392

QY 444 SGSVGGTYSKPHHTMQVCTONVNTIDEIPPENE--PLSRGYSRLSHITSY--SPSKNA 499
DB 393 SSNLPASVYRKSGTV-----DSLDEIPQNNVPPRQGFSLSHSVMSFRSGFSNSS 444

QY 500 SSPARYGNLVPFAWTHRSADVNTVYSKIQIOPVVKAAHTLVSGTIVYKGPFTGNGILK 559

A:Accession: JC2219
A:Molecule type: DNA
A:Residues: 1-1176 <UDA>
A:Cross-references: DDBJ:D17518; NID:g506190; PIDN:BAA04468.1; PID:g535781
C:Genetics:
A:Gene: cryIA(a)
C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 33.8%; Score 2045; DB 2; Length 1176;
Best Local Similarity 39.3%; Pred. No. 1.9e-110;
Matches 479; Conservative 174; Mismatches 415; Indels 152; Gaps 26;

Qy 53 NPDL-----FCNPET-FISSSTVQTG---IGI-----VGVGLGALGVPPAGQIASFYSP 97
Db 4 NPINECIPYCNLSNPEVEVLGGERIETGTPIDISLSLTQPLSEFPVPGAGFVLGLVDI 63
Qy 98 IVGQLWPSSVSVWEMIMKQVEDLIDQKI TDSVRKTALAGLQGLGDLDVYQKSLKNWLE 157
Db 64 IWGIFGPSQ----WDAFLVQIEQLISQRIEFPARNQAISRLGLSNLYQIYAESFWEA 119
Qy 158 NRNDTRARSVVVYQYIALELDFVAKIPSAISGQVEPLLSVYQAANLHLLRLDASIFG 217
Db 120 DPTNPALREMRIOFNDMSALTTAIFLLAVQNVQVPELLSVYQVQANLHLSVLDRVSFG 179
Qy 218 AEWGFTPGEISTFYDRQVTTAQSDYCVKMYNTGLDKLKTNAASMLKYHQFRREMTLL 277
Db 180 QRWGLDVATINSRYNDLTRIGTYTDAVRMYNTGLERVWGPDSRDWVYQFRBELTIT 239
Qy 278 VLDLVALFPNYDTRTYPIETTAQLTRVYTDPIVFNRETSGGPCRRWSLNSDISFSEVES 337
Db 240 VLDIVSLFPNYSRTYPIRTVSQLTRTYTPVLETFD--GSFRG-----SAQRIEQ 289
Qy 338 AVTRSPHLFDILSIEFTTTRAGLPLANTYEZLEWVGHSHK-----YKNTWASSALERNY 392
Db 290 S-IRSPHMLDILNSITTYTDAHG-----GYTWSHGQIMASVPVGGPBPFPPL---Y 338
Qy 393 GTTISNKIKYDYLANKDIFQVRSILGADLANYAQ--VYGVVPASFTLLDKNT-----GSGS 446

913 ATLGNELVEEGPLSGETLERAQQEQOQWODKMKRKGSEKAYAAKOALDRLPADYOD 972
857 ARLGNELEFLSEKPLVGEARLAKVRAEKMKRKELEWETNIVYKRAKESVDALFVRSQY 916
973 QKLSNGVEMSDMLAAQVLVQSIPIVYVNDALPEIPGMYNTTSFTELTNLQQAWNLVDLNA 1032
917 DQLQADTNIAIHAADEKRVHSIREAYLPVLPVGNAAIFEELEGRISTAFSLYDARNV 976
1033 IPNGDFRNLGSLDNWATSDVNV--QOLSDTSLVIVPNNWSQVSQOFTVQPNRYVLRVATRK 1091
977 IKNGDFNGLSCWNVKGVHDVEBQNNQSRSLVVPWEAREVSOEVRVCPGRGYILLRVATYK 1036
1092 EGVGDGVVLIIRDCANQNTETLTEN-----ICDD----- 1118
1037 EGYGEGCVTHIETNMDFELKNCVBEIYPNNTVTCNDYTVNQBEYVGAYTSNRGYN 1096
1119 DTVGLSADOTS-----YITKTVEFTPTPTEQVWID 1147
1097 EAPSPADYASVYBEKSYTDGRRNPCEFNRGYDYTPLPVGVYTKLELYPPEPDKVWIE 1156
1148 MSETGVENIESVELVLEHE 1167
1157 IGETEGTFIVDSVELLIMEE 1176

RESULT 14
A41052
Parasporal crystal protein cryAel - Bacillus thuringiensis (strain alesti)
C:Species: Bacillus thuringiensis
C:Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 01-Dec-2000
C:Accession: A41052
R:Lee, C.S.; Aronson, A.I.
J. Bacteriol. 173, 6635-6638, 1991
A:Title: Cloning and analysis of delta-endotoxin genes from Bacillus thuringiensis subsp.
A:Reference number: A41052; MUID:92011442; PMID:1655719
A:Accession: A41052
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1181 <LEE>
A:Cross-references: GB:M65252; NID:g142874; PIDN:AAA22410.1; PID:g142875
C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 33.8%; Score 2044.5; DB 2; Length 1181;
Best Local Similarity 39.0%; Pred. No. 1.9e-110;
Matches 477; Conservative 190; Mismatches 404; Indels 151; Gaps 29;

Qy 53 NPDL-----FCNPET-FISSSTVQTG---IGI-----VGVGLGALGVPPAGQIASFYSP 97
Db 4 NPINECIPYCNLSNPEVEVLGGERIETGTPIDISLSLTQPLSEFPVPGAGFVLGLIDL 63
Qy 98 IVGQLWPSSVSVWEMIMKQVEDLIDQKI TDSVRKTALAGLQGLGDLDVYQKSLKNWLE 157
Db 64 IWGIFGPSQ----WDAFLVQIEQLISQRIEFPARNQAISRLGLSNLYQIYAEAPREWA 119
Qy 158 NRNDTRARSVVVYQYIALELDFVAKIPSAISGQVEPLLSVYQAANLHLLRLDASIFG 217
Db 120 DPTNPALREMRIOFNDMSALTTAIFLTQVNTQVPLLSVYQVQANLHLSVLDRVSFG 179
Qy 218 AEWGFTPGEISTFYDRQVTTAQSDYCVKMYNTGLDKLKTNAASMLKYHQFRREMTLL 277
Db 180 QRWGLDVATINSRYNDLTRIGTYTDAVRMYNTGLERVWGPDSRDWVYQFRBELTIT 239
Qy 278 VLDLVALFPNYDTRTYPIETTAQLTRVYTDPIVFNRETSGGPCRRWSLNSDISFSEVES 337
Db 240 VLDIVSLFPNYSRTYPIRTVSQLTRTYTPVLETFD--GSFRG-----SAQRIEQ 289
Qy 338 AVTRSPHLFDILSIEFTTTRAGLPLANTYEZLEWVGHSHK-----YKNTWASSALERNY 392
Db 290 S-IRSPHMLDILNSITTYTDAHG-----GYTWSHGQIMASVPVGGPBPFPPL---Y 338
Qy 393 GTTISNKIKYDYLANKDIFQVRSILGADLANYAQ--VYGVVPASFTLLDKNT-----GSGS 446

```
Db 339 GTMGNAAPQQRIVAGLGGQVVRITLS---STYRNPFFIIGINNQLSLVDLGTPEYAGSSSN 395
Qy 447 VGGTYSKPHHTMQVCTQNYATIDIEIPPE--NEPLSRGYSHRLSHITSY--SPSKNASSP 502
Db 396 LPSAVYRKSGTV-----DSLDBIIPQDNNVPFPGQSHRLSHVSNFRSGFSSSVSI 447
Qy 503 ARYGNLPVPWATHRSADVNTVYSDKJTOIPVKAHTLVSGTTVVKGPFTGGNILKRTS 562
Db 448 IR---APMPSWIHRSABFNNIIPSSQITQIPLTSTNLGSGTSVVGKPGFTGGDLRRTS 504
Qy 563 SGPLAYSVSVKSPLSQRYRARIYASTTNLRLFTVTSIGTRIYSINVNKTKMKGDDLTEN 622
Db 505 PGQISTLRVNTAPLSQRYRARIYASTTNLQFTHTSIDGRPNQGNFSATMSSGNGLSQ 564
Qy 623 TFDLATIGTAPTPSNYSDSLTVGADSPASGEVYVDKPELIPVNAATFEABEDLDVAKKAV 682
Db 565 SFRVTGFTTPNFNGSGSVFTLSAHVFNSGNEVYIDRIEFVPAEVTFEAYDLERAQEA 624
Qy 683 NGLFTSKKD-ALQTSVTDYQVNOAANLVECLSDBELYPNEKMLMDAVKEAKRLVQARNLL 741
Db 625 NALFTSPNQIGLTKDVTDYHIDQVSNLVECLSDBELYKRELSKVKHAKRLSDERNLL 684
Qy 742 QDTCFNRLN--GENGWTSSTGIEVABGDVLPKDRSLRLTSAREIDTPTPTLYQOQIDES 799
Db 685 QDPNFRGINQPDGRWGRGTDITIQGGDDVFKENYVILPGTFD---ECYPTLYQKIDES 741
Qy 800 LLKPYTRYKLGKFTIGSSODLEIKLIRHRANOIVKNVP--DNLLPDVL--PVNSCGGIDRC 855
Db 742 KLKAYTRYELRGVIEDSQDLIELYIRYNAKHETVNVPGTGLWPLSPFESSIGKCEPNRC 801
Qy 856 SEQOYVDANLALENNGENG--NMSDSSHAPSPIHDTGHEIDNENTGIWVVFVKIPPTNGVAT 914
Db 802 A--PHLEWNPDLDCSQRDGEKCAHSHHPSLIDIDVGCIDILNEDLGVMVIFKIKTQDGHAR 859
Qy 915 LGNLELVEEGSLGETLERAAQQOQWQDKMARKGSEKAYYAKOALDELFLADYQDQK 974
Db 860 LGNLEFLBEKPLVGEALARVRAKKWRDKREKLEQLETNIVYKEAKESVDALFVNSQYDQ 919
Qy 975 LNSGVNSDMLAQNVLQSIPIVYVNDALPETIPGNVYTSFTELNRLOQAWMLYDLRNLAP 1034
Db 920 LQADTNIAHTADKRVHRIQEAYLPELSPVPGVAGIFEELEGRIPTAYSILYDARNVVK 979
Qy 1035 NGDFRNLGDWATSDVNV--QOQSDTSVLVTPNMSQVSOQFTQPNRYVYLRVTKARKEG 1093
Db 980 NGDFNGLSLCNVAGHVDVEQNNHRSVLYVPEWEAEVSQEVRCVPCRGYTLRVYATKEG 1039
Qy 1094 VGDGVLIIRDCANOTETLTENIC-----DDG 1119
Db 1040 YGEGCVTIHETENNIDELKFCNVEEVYFNNTVTCNEYTANQEEYGGAYTSCNRGYDET 1099
Qy 1120 TG---VLSADQTS-----YITKTVFETPSTEQVW 1145
Db 1100 YGSNYSPADYASYEKEAYTDGRRENPCESNRGYDGYTLPAGVTKQLEYFPETDKVW 1159
Qy 1146 IDMSSETGVFNIRSVELVLEER 1167
Db 1160 IEIGETEGTFTVDSVELFLMER 1181

RESULT 15
A29125
paraasporal crystal protein Bc2 - Bacillus thuringiensis subsp. kurstaki (strain HD-1)
C:Species: Bacillus thuringiensis subsp. kurstaki
C:Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 01-Dec-2000
C:Accession: A29125
R:Fischhoff, D.A.; Bowdish, K.S.; Perlak, F.J.; Marrone, P.G.; McCormick, S.M.; Niederme
Bio/Technology 5, 807-813, 1987
A:Title: Insect tolerant transgenic tomato plants.
A:Reference number: A29125
A:Accession: A29125
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-1156 <PIS>
```

C:Superfamily: paraasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 33.8%; Score 2043; DB 2; Length 1156;
Best Local Similarity 39.1%; Pred. No. 2.2e-110;
Matches 476; Conservative 185; Mismatches 389; Indels 168; Gaps 30;

```
Qy 53 NPEL-----FQNPET-FISSSTVQTG-----IGI-----VGQVLGALGVPPAGQIASPVSF 97
Db 4 NPNINECIPYCNLENPEVEVLGGRITGYTPIDISLSTQTLFLLSBFPVGGAFVLGLVDI 63
Qy 98 IVGQLWPSSTVSVMEMKQVEDLIQDKITDSVKRTALAGLGGDGLDVYQKSLKNWLE 157
Db 64 IWGIFGSPQ-----WDAFLVQIEQINRIEFARQALSRLEGLSNLYQIYASSPREHEA 119
Qy 158 NRNDTRARSVVVTVQYIALELDFVAKIPSAFISGOEVPLLSVYQAANLHLALLLRLDASIFG 217
Db 120 DPTNPALREEMRIQFNDMNSALTTAIPLFAVQYQVPLLSVYQAANLHLSVLRDVSFVG 179
Qy 218 AEWGFTPGEISTFDYDQVTRTAQSYDVCVKQYNTGLDKLKGNTNAASWLKYHQFREMILL 277
Db 180 QRWGFDAATINSRYNDLTRLIGNYTDHAVRYNTGLERWGPDSRDMTRYNQFRELTLT 239
Qy 278 VLDLVALFPNVDYETPTPIETTAQLTRVYTDPIVFNRETSGFCRRWGLNSD1SPSEVES 337
Db 240 VLDIVSLFENYDSKTYPIRTVSQLTREIYNPVLNFD--GSP--RGSAGQ-----LEG 289
Qy 338 AVIRSPHLFDILSIEFIYTRAGLPLANTLEYLWVGHSHK-----YKNTNASSALERNY 392
Db 290 S-IRSPHLMDILNSITIVT-----DAHGEY--YWSGHQIMASPVGSPGPBPTPL--Y 338
Qy 393 GTITSNKKIYYDLANKDIPQVRSIGADLANIYAQVY--GVPIYAFLLDKVT-----GSGS 446
Db 339 GTMGNAAPQQRIVAGLGGQVVRITLSSTL---YRPPFNIIGINNQLSLVDLGTPEYAGTSSN 395
Qy 447 VGGTYSKPHHTMQVCTQNYATIDIEIPPE--NEPLSRGYSHRLSHITSY--SPSKNASSP 502
Db 396 LPSAVYRKSGTV-----DSLDBIIPQDNNVPFPGQSHRLSHVSNFRSGFSSSVSI 447
Qy 503 ARYGNLPVPWATHRSADVNTVYSDKJTOIPVKAHTLVSGTTVVKGPFTGGNILKRTS 562
Db 448 IR---APMPSWIHRSABFNNIIPSSQITQIPLTSTNLGSGTSVVGKPGFTGGDLRRTS 504
Qy 563 SGPLAYSVSVKSPLSQRYRARIYASTTNLRLFTVTSIGTRIYSINVNKTKMKGDDLTEN 622
Db 505 PGQISTLRVNTAPLSQRYRARIYASTTNLQFTHTSIDGRPNQGNFSATMSSGNGLSQ 564
Qy 623 TFDLATIGTAPTPSNYSDSLTVGADSPASGEVYVDKPELIPVNAATFEABEDLDVAKKAV 682
Db 565 SFRVTGFTTPNFNGSGSVFTLSAHVFNSGNEVYIDRIEFVPAEVTFEAYDLERAQEA 624
Qy 683 NGLFTSKKD-ALQTSVTDYQVNOAANLVECLSDBELYPNEKMLMDAVKEAKRLVQARNLL 741
Db 625 NALFTSPNQIGLTKDVTDYHIDQVSNLVECLSDBELYKRELSKVKHAKRLSDERNLL 684
Qy 742 QDTCFNRLN--GENGWTSSTGIEVABGDVLPKDRSLRLTSAREIDTPTPTLYQOQIDES 799
Db 685 QDPNFRGINQPDGRWGRGTDITIQGGDDVFKENYVILPGTFD---ECYPTLYQKIDES 741
Qy 800 LLKPYTRYKLGKFTIGSSODLEIKLIRHRANOIVKNVP--DNLLPDVL--PVNSCGGIDRC 855
Db 742 KLKAYTRYELRGVIEDSQDLIELYIRYNAKHETVNVPGTGLWPLSPFESSIGKCEPNRC 801
Qy 856 SEQOYVDANLALENNGENG--NMSDSSHAPSPIHDTGHEIDNENTGIWVVFVKIPPTNGVAT 914
Db 802 A--PHLEWNPDLDCSQRDGEKCAHSHHPSLIDIDVGCIDILNEDLGVMVIFKIKTQDGHAR 859
Qy 915 LGNLELVEEGSLGETLERAAQQOQWQDKMARKGSEKAYYAKOALDELFLADYQDQK 974
Db 860 LGNLEFLBEKPLVGEALARVRAKKWRDKREKLEQLETNIVYKEAKESVDALFVNSQYDQ 919
Qy 975 LNSGVNSDMLAQNVLQSIPIVYVNDALPETIPGNVYTSFTELNRLOQAWMLYDLRNLAP 1034
Db 920 LQADTNIAHTADKRVHRIQEAYLPELSPVPGVAGIFEELEGRIPTAYSILYDARNVVK 979
Qy 1035 NGDFRNLGDWATSDVNV--QOQSDTSVLVTPNMSQVSOQFTQPNRYVYLRVTKARKEG 1093
Db 980 NGDFNGLSLCNVAGHVDVEQNNHRSVLYVPEWEAEVSQEVRCVPCRGYTLRVYATKEG 1039
Qy 1094 VGDGVLIIRDCANOTETLTENIC-----DDG 1119
Db 1040 YGEGCVTIHETENNIDELKFCNVEEVYFNNTVTCNEYTANQEEYGGAYTSCNRGYDET 1099
Qy 1120 TG---VLSADQTS-----YITKTVFETPSTEQVW 1145
Db 1100 YGSNYSPADYASYEKEAYTDGRRENPCESNRGYDGYTLPAGVTKQLEYFPETDKVW 1159
Qy 1146 IDMSSETGVFNIRSVELVLEER 1167
Db 1160 IEIGETEGTFTVDSVELFLMER 1181
```

Db 899 TNIAITHAADRVHSIREAYLPBLSVIPGVNAALFBELEGRIPTAFSLYDARNVKNKNDP 958
Qy 1039 RNLSDNATSDNVV-QQLSDTSVLVIPNMSQVSQQFTVQPNRYVYLVRTARKEGVGDG 1097
Db 959 NGLSCHNVKGVHVDVEQNNHRSVLVWVPEAEVSVQVRVCPGRGYILRVTVAYKEGYGEG 1018
Qy 1098 YVIRDCANQTEITLTFN-----ICDDDTGV----- 1122
Db 1019 CVTIHEIENNTDELKFSNCVBEVYPNNVTVCNDYTATQBEYEGTYSRNRGYDAYS 1078
Qy 1123 --LSADQTS-----YITKVFTPSTEQVWIDMS 1149
Db 1079 SSVPADYASAYEERKAYTDGRDNPCESNRGCHDYTPLPAGYVTKELYPPETDKWIEIG 1138
Qy 1150 ETEGVFNIESVELVLEE 1167
Db 1139 ETEGTFIVDSVELLMEE 1156

Search completed: June 21, 2004, 13:39:44
Job time : 34 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	4007	66.3	1157	1	C8AA_BACTK	Q45704 bacillus th
2	3733.5	61.8	1160	1	C8CA_BACTP	Q45706 bacillus th
3	3152	52.2	1169	1	C8BA_BACTK	Q45705 bacillus th
4	3123	51.7	1157	1	C9CA_BACTO	Q45733 bacillus th
5	2825.5	46.7	1150	1	C9EA_BACTA	Q9zn19 bacillus th
6	2715	44.9	1169	1	C9DA_BACTP	Q06014 bacillus th
7	2714	44.9	1215	1	C1KA_BACTM	Q45715 bacillus th
8	2552	42.2	1229	1	C1BB_BACTU	Q45739 bacillus th
9	2547.5	42.1	1238	1	C1BA_BACTK	P05517 bacillus th
10	2491	41.2	1233	1	C1BC_BACTM	Q45774 bacillus th
11	2462	40.7	1227	1	C1BE_BACTU	O85805 bacillus th
12	2458.5	40.7	1138	1	C7AA_BACTU	Q03749 bacillus th
13	2458.5	40.7	1138	1	C7AB_BACTA	Q45707 bacillus th
14	2445	40.5	1231	1	C1BD_BACTZ	Q9zaz5 bacillus th
15	2439.5	40.4	1138	1	C7AB_BACTK	Q45708 bacillus th
16	2327.5	36.9	1163	1	C0AA_BACTP	Q9z597 bacillus th
17	2191	36.3	1156	1	C9AA_BACTG	Q99031 bacillus th
18	2154.5	35.6	1174	1	C1FA_BACTA	Q03746 bacillus th
19	2093.5	34.6	1166	1	C1GA_BACTU	Q05746 bacillus th
20	2065.5	34.2	1167	1	C1JA_BACTU	Q45738 bacillus th
21	2065	34.2	1170	1	C1JB_BACTU	Q45716 bacillus th
22	2061.5	34.1	1155	1	C1AB_BACTK	P06578 bacillus th
23	2044.5	33.8	1181	1	C1AE_BACTL	Q03748 bacillus th
24	2038	33.7	1176	1	C1AA_BACTK	P02965 bacillus th
25	2024.5	33.5	1171	1	C1EA_BACTX	Q57458 bacillus th
26	2011.5	33.3	1179	1	C1AD_BACTA	Q03744 bacillus th
27	1999	33.1	1169	1	C1GB_BACTZ	Q9zaz6 bacillus th
28	1995	33.0	1169	1	C1CA_BACTE	P05518 bacillus th
29	1991	32.9	1169	1	C1PB_BACTM	O66377 bacillus th
30	1953.5	32.3	1160	1	C1DB_BACTU	Q45747 bacillus th
31	1946.5	32.2	1174	1	C1EB_BACTA	Q03745 bacillus th
32	1944	32.2	1165	1	C1DA_BACTA	P19415 bacillus th
33	1939	32.1	1178	1	C1AC_BACTK	P05068 bacillus th

Db 1 MSPNNQNEYIIDATPSTSVSSDSNRYPPFANEPTDALQNNYKDYLMQSGGGENPELFCNP 60
Qy 61 ETFISSSTVQTGIGIVGVLGALGVPFAGQIASFYSEIVGLQWPSSTSVVMEMIMKQVED 120
Db 61 ETFISSSTVQTGIGIVGVLGALGVPFAGQIASFYSEIVGLQWPSSTSVVMEMIMKQVED 120
Qy 121 LIDOKITSDVKTALAGLOGDGLDVYQKSLKWLNRNDRTRARSVVVYQYIALELDV 180
Db 121 LVDOKIERYKDKALAEKGLGNALDVYQKSLKWLNRNDRTRARSVVVYQYIALELDV 180
Qy 181 AKIPSAISGEVPLLSVYAAQANLHLLLDASIFGAEMGFTPCGEISTFYVDROVTTAQ 240
Db 181 SSISFPAVSGHEVLLVAYAVNLHLLLDASIFGEEMGFTPCGEISTFYVDROVTTAQ 240
Qy 241 YSDYCVMYNTGLDKLGTNAASLKYHQFPRMTLLVLDLVALFPNPDYRTYETTTAQ 300
Db 241 YSDYCVMYKGLDKLGTTSKMLNYHQFPRMTLLVLDLVALFPNPDYRTYETTTAQ 300
Qy 301 LTRVYVTDPIVNRETSGGFCRRVSLNSDISFSEVESAVIRSPHLFDILSIEFTYTRAG 360
Db 301 LTRVYVTDPIAFNIVTSTGFCNPWSTHSGILFYEVNNVIRPPELFDILSIEFTYTRAG 360
Qy 361 LPLANTEVLEWVWCHSTKYKNT-NASSALERNYGTITSNKIKYYDLANKDIFOVRSIGAD 419
Db 361 ITLNDAYINWGHTRKRTADSTVITYANYGRITSEK-NSPALEDORDIFEINSTVAN 419
Qy 420 LANYAQQVYGVYFASFTLLDKNTGSGVGGFTYSKPHFTMQVCTQNTYDIEIPPENE-P 478
Db 420 LANYQKAYGVGFWFHWK-RTSSTAYLYSKTHALOGCTQVYESSDEIPLDRVP 477
Qy 479 LSRGSHRLSHITSYSPSKNASSPARVGNLVPFANTHRASDVNTVTSYDCKITQIPVKAH 538
Db 478 VABSYSHRLSHITSYSPSKNASSPARVGNLVPFANTHRASDVNTVTSYDCKITQIPVKAH 535
Qy 539 TLVSGTVLKGPGTGNILARTSGGLAYTSVSVKSPLSQRYRARIYASTNLRFLVT 598
Db 536 MLXGGSVVQPGFTGGDILKRTPSILGTFAVTVNGSLQRYRARIYASTNLRFLVT 594
Qy 599 ISGTRIYSINVKTMKDDITFTNFDTLATITGTAFTFNSYDSLTIVGADSPASGEVYVD 658
Db 595 YLGDTIBNRNPKMTDNGASLYETFPKASFTDPQPRETQDKILLSNGDFSQQEVYID 654
Qy 659 KPELIPVNAFEABEDLDVAKAVNGLFTSKDALQTSVTDYQVNOANLVECLSDLYP 718
Db 655 RIEFIPVDYEAQDLEAAKAVNALFTNTKQGLRPGVTDYEVNOANLVECLSDLYP 714
Qy 719 NEKMLMDAVKELVQARNLQDTPGNRNGENGWGTGIEVAEGDVLFPDORSRLT 778
Db 715 NEKMLMDAVKELVQARNLQDTPGNRNGENGWGTGIEVAEGDVLFPDORSRLT 774
Qy 779 SAREIDTETPTLYQQIDESLLKPYTRYKLGFTGSSQDLEIKLIRANQIVKVNPDN 838
Db 775 GAREIDTETPTLYQVKEGVLPKPYTRYLRGFPVGSQGLEIYVIRHQTNRVKNVDD 834
Qy 839 LLPVLVWNSCGGIDRCSEQQYVDANLALNNGENGWSSSHAPSHIDGCEIDLMENT 898
Db 835 LLPVSPVNSDGSINRCSEQYVNSRL-----EGENRSGDAHEFSLPIDIGELDYNENA 888
Qy 899 GIWVVPKLPITNGVATLGNLVEBGLSGETLERAQOQOQOQOQKMARKGASAKAYA 958
Db 889 GIWVGFKLIDEGVATLGNLVEBGLSGDALERLQREBQOMKQTRRREEDDRRTMA 948
Qy 959 AKQADRLFYDQOKLNGSVKSMALAAQNLVQSIPIVYNDALPEIPGMNYTSTFELTN 1018
Db 949 SKQAVRLYADYQDQQLNPVEITDLTAQDLIQSIPVYVNMPEPBEIPGMNYTSTFELTD 1008
Qy 1019 RLQQAANLYDURNALPNGDFRNLGSDMNATSDNVVQQLSDTSVLVIPNWSQVSOQFTVQ 1078
Db 1009 RLQQAANLYDQRNAIPNGDFRNLGSDMNATPGVEVQQINHTSVLVIPNWSQVSOQFTVQ 1068
Qy 1079 PNYRVVLTARKEGSGVGYIIRGANQOTETLTFTNICDDDTGLVSDQTSYITKTVFT 1138
Db 1069 PNQRYVLTARKEGSGVGYIIRGANQOTETLTFTNICDDDTGLVSDQTSYITKTVFT 1128

Qy 1139 PSTQVWIDMSSETGVFNIESVELVLBER 1167
Db 1129 PYTDQWIEISETGTFVIESVELIVDVE 1157

RESULT 2

C8CA BACTP
ID C8CA BACTP STANDARD; PRT; 1160 AA.
AC Q45706;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Pesticidal crystal protein crysca (insecticidal delta-endotoxin
DE CRYVIIIC(a) (crystalline entomocidal protoxin) (130 kDa crystal
DE protein).
GN CRY8CA OR CRYVIIIC(A) OR CRYIIII.
OS Bacillus thuringiensis (subsp. japonensis).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=128936;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Buibui.
RX MEDLINE=94100786; PubMed=7764305;
RA Sato R., Takeuchi K., Ogiwara K., Minami M., Kaji Y., Suzuki N.,
Hori H., Asano S., Ohba M., Iwahana H.,
RT "Cloning, heterologous expression, and localization of a novel crystal
RT protein gene from Bacillus thuringiensis serovar japonensis strain
RT buibui toxic to scarabaeid insects.";
RL Curr. Microbiol. 28:15-19(1994).
RN (2)
RP SEQUENCE OF 1-14 AND 56-64, AND CHARACTERIZATION.
RX MEDLINE=94259659; PubMed=8200856;
RA Hori H., Suzuki N., Ogiwara K., Himejima M., Indrasith L.S.,
RA Minami M., Asano S., Sato R., Ohba M., Iwahana H.,
RT "Characterization of larvicidal toxin protein from Bacillus
RT thuringiensis serovar japonensis strain Buibui specific for
RT scarabaeid beetles.";
RL J. Appl. Bacteriol. 76:307-313(1994).
CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF INSECTS. ACTIVE ON VARIOUS SCARABAEID BEETLES
CC SUCH AS ANOMALA CUPREA, A. RUFOCUPREA AND POPILLIA JAPONICA.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U04366; AAA21119.1; -.
DR PIR; I40589; I40589.
DR HSP; P07130; 1DLC.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal bind_like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
KW Toxin; Sporulation.
SQ SEQUENCE 1160 AA; 130425 MW; C16C3D912ESB8751 CRC64;

Query Match 61.8%; Score 3733.5; DB 1; Length 1160;
Best Local Similarity 64.2%; Pred. No. 4.4e-205;
Matches 767; Conservative 109; Mismatches 257; Indels 61; Gaps 21;

```
QY 1 MSPNNQNEYIILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSGEGNPELPGNP 60
DB 1 MSPNNQNEYIILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSGEGNPELPGNP 60
QY 61 EFTFIS-STVQTGIGIVGVLGALGVPFAGQIASFYSPVIGOLWPSSTVSUWEMIKQVE 119
DB 61 GTFISAQDAVGTDIVSTIISGLGIPVLGVEFSILGSLGLLWPSNNENWQIIPNVRVE 120
QY 120 DLIDQKITDSVRKTAALAGLQGLGDVYQKSLKNWLENRNDTRARSVVVTVQYIALELDF 179
DB 120 DLIDQKITDSVRKTAALAGLQGLGDVYQKSLKNWLENRNDTRARSVVVTVQYIALELDF 179
QY 121 ELIDOKILDSVRRAIDLANSLRIAVFYQNALEDWFKNPHSTRNALVKRPGNAEAIL 180
DB 121 ELIDOKILDSVRRAIDLANSLRIAVFYQNALEDWFKNPHSTRNALVKRPGNAEAIL 180
QY 180 VAKIPSPAISGOVPLLSVYQAANLHLLLRDASIFGABWGFPGTGEISTFYDROVTRTA 239
DB 180 VAKIPSPAISGOVPLLSVYQAANLHLLLRDASIFGABWGFPGTGEISTFYDROVTRTA 239
QY 181 RTNMGSPSTNYETETPLLPYQAASLHLVMDVDQVIGKEMWGPQNDIDLKYKEQVSUYTA 240
DB 181 RTNMGSPSTNYETETPLLPYQAASLHLVMDVDQVIGKEMWGPQNDIDLKYKEQVSUYTA 240
QY 240 QYSDYCVKQWYNTGLDKLGTNAASHLKHQPREMLTLLVLDVALPNNVDITYPLETIA 299
DB 240 QYSDYCVKQWYNTGLDKLGTNAASHLKHQPREMLTLLVLDVALPNNVDITYPLETIA 299
QY 241 RYSDHCQWYNAGLNKLRGTGAKQWVDVNRFRFEMNVMVLDVALPNNVDIARIYPLETNA 300
DB 241 RYSDHCQWYNAGLNKLRGTGAKQWVDVNRFRFEMNVMVLDVALPNNVDIARIYPLETNA 300
QY 300 OLTRVYTDPIVFNRETSGGFCRWSLN-----SDI-----SFSEVESAVIRSPHLEDIL 349
DB 300 OLTRVYTDPIVFNRETSGGFCRWSLN-----SDI-----SFSEVESAVIRSPHLEDIL 349
QY 301 ELTREIFIDPV-----GSYVTGSSSTLISWYDMIPALPSPSTLEN-LARKPDFFTL 352
DB 301 ELTREIFIDPV-----GSYVTGSSSTLISWYDMIPALPSPSTLEN-LARKPDFFTL 352
QY 350 SBIETPTTRAGLPLNNT-EYLEYVWGHSLIKYKVTNASSALERNYGTITSNKIKYYDLANK 408
DB 350 SBIETPTTRAGLPLNNT-EYLEYVWGHSLIKYKVTNASSALERNYGTITSNKIKYYDLANK 408
QY 353 QBIRWMTS---PRQMTIETIYNYWQGRLTLSYVGGSP--NKYSGLVLAGAEDIIIPVQGN 407
DB 353 QBIRWMTS---PRQMTIETIYNYWQGRLTLSYVGGSP--NKYSGLVLAGAEDIIIPVQGN 407
QY 409 DIFQVRSIGADLANTYAQVGYVPYASFLLDIDWGTSGSGVGGTYSKPHMTMQVCTQNTY 468
DB 409 DIFQVRSIGADLANTYAQVGYVPYASFLLDIDWGTSGSGVGGTYSKPHMTMQVCTQNTY 468
QY 408 DIYRV--VMTYIGRTYNSLLGNVPYTF-YFSNNYTK-----TYSKP-----KOFAGGIKT 454
DB 408 DIYRV--VMTYIGRTYNSLLGNVPYTF-YFSNNYTK-----TYSKP-----KOFAGGIKT 454
QY 469 ID---EIPPENPLSGYSHRLSHITSYFSKNASPARYGNLPPAWTHRSADVTNTVY 525
DB 469 ID---EIPPENPLSGYSHRLSHITSYFSKNASPARYGNLPPAWTHRSADVTNTVY 525
QY 455 IDSGBELTYEN---VQSYSHRVSYSITSFEIKSTGTV--LGVVPFGWTHSSASRNNFY 509
DB 455 IDSGBELTYEN---VQSYSHRVSYSITSFEIKSTGTV--LGVVPFGWTHSSASRNNFY 509
QY 526 SDKITQIPVVKATLVSGTIVTKGPG-FYTGNIKRTSSGPLAY---TSVSVKSPLSQRY 581
DB 526 SDKITQIPVVKATLVSGTIVTKGPG-FYTGNIKRTSSGPLAY---TSVSVKSPLSQRY 581
QY 510 ATKISQIPINKASRTSGGAVNPFQGLYNGPVMKLSGSGSVINLRVATDAKG-ASQRY 568
DB 510 ATKISQIPINKASRTSGGAVNPFQGLYNGPVMKLSGSGSVINLRVATDAKG-ASQRY 568
QY 582 RARIYASTNLRFLWTIS-----GTFIYINVNKTANKGDDLTFTNFDLATIGTA-F 633
DB 582 RARIYASTNLRFLWTIS-----GTFIYINVNKTANKGDDLTFTNFDLATIGTA-F 633
QY 569 RIRIYASDRAGK--TISSRSPENPATYSASIAVTNTMTNASTYTSFYAEGSPINL 626
DB 569 RIRIYASDRAGK--TISSRSPENPATYSASIAVTNTMTNASTYTSFYAEGSPINL 626
QY 634 TFSNYSDSLTVGADSPASGGEVYVDKFLIPVNATFEAEDLDVAKAVNGLFTSKDAL 693
DB 634 TFSNYSDSLTVGADSPASGGEVYVDKFLIPVNATFEAEDLDVAKAVNGLFTSKDAL 693
QY 627 GISGSRPTDISITKBAAGNLIYDRIEPIPVNTLFEAEDLDVAKAVNGLFTNEKDAL 686
DB 627 GISGSRPTDISITKBAAGNLIYDRIEPIPVNTLFEAEDLDVAKAVNGLFTNEKDAL 686
QY 694 QTSVTDYQVNOANLVECLSDLYPNEKRMMLDVAKEAKRLVOARNLLQDTGFNRINGEN 753
DB 694 QTSVTDYQVNOANLVECLSDLYPNEKRMMLDVAKEAKRLVOARNLLQDTGFNRINGEN 753
QY 687 QTSVTDYQVNOANLVECLSDLYPNEKRMMLDVAKEAKRLVOARNLLQDTGFNRINGEN 746
DB 687 QTSVTDYQVNOANLVECLSDLYPNEKRMMLDVAKEAKRLVOARNLLQDTGFNRINGEN 746
QY 754 GWTGTGTIEVAGDVLFKDRSLRLTSAREIDTETPTLYLQIDSLKPKYTRYLKGFI 813
DB 754 GWTGTGTIEVAGDVLFKDRSLRLTSAREIDTETPTLYLQIDSLKPKYTRYLKGFI 813
QY 747 GWTGTGTIEVAGDVLFKDRSLRLTSAREIDTETPTLYLQIDSLKPKYTRYLKGFI 806
DB 747 GWTGTGTIEVAGDVLFKDRSLRLTSAREIDTETPTLYLQIDSLKPKYTRYLKGFI 806
QY 814 GSSQDLKILKIRHRANOIVKNVPMNLLPDVLPVNSCGGIDRCSEQYVDANLALENNEN 873
DB 814 GSSQDLKILKIRHRANOIVKNVPMNLLPDVLPVNSCGGIDRCSEQYVDANLALENNEN 873
QY 807 GSSQDLKILKIRHRANOIVKNVPMNLLPDVLPVNSCGGIDRCSEQYVDANLALENNEN 866
DB 807 GSSQDLKILKIRHRANOIVKNVPMNLLPDVLPVNSCGGIDRCSEQYVDANLALENNEN 866
QY 874 GNMSSDSHAFPHIDTGEIDLNTGNTGIWVVKIPPTNGYATIGNLVEBGLSGETLER 933
DB 874 GNMSSDSHAFPHIDTGEIDLNTGNTGIWVVKIPPTNGYATIGNLVEBGLSGETLER 933
QY 867 GNMSSDSHAFPHIDTGEIDLNTGNTGIWVVKIPPTNGYATIGNLVEBGLSGETLER 926
DB 867 GNMSSDSHAFPHIDTGEIDLNTGNTGIWVVKIPPTNGYATIGNLVEBGLSGETLER 926
QY 934 AQOQSQWODKWKARGASEKAYAAKQAI DRLFADYQDQKLSNGVMSDMLAAQNLVQS 993
DB 934 AQOQSQWODKWKARGASEKAYAAKQAI DRLFADYQDQKLSNGVMSDMLAAQNLVQS 993
QY 927 AQOQSQWODKWKARGASEKAYAAKQAI DRLFADYQDQKLSNGVMSDMLAAQNLVQS 986
DB 927 AQOQSQWODKWKARGASEKAYAAKQAI DRLFADYQDQKLSNGVMSDMLAAQNLVQS 986
QY 994 IPYVNDALPEIPGMNYTSPTELNRLOQAWNLXDLRNAI PNGDPFNGLSDMNATSDVNV 1053
DB 994 IPYVNDALPEIPGMNYTSPTELNRLOQAWNLXDLRNAI PNGDPFNGLSDMNATSDVNV 1053
QY 987 IPYVNDALPEIPGMNYTSPTELNRLOQAWNLXDLRNAI PNGDPFNGLSDMNATSDVNV 1046
DB 987 IPYVNDALPEIPGMNYTSPTELNRLOQAWNLXDLRNAI PNGDPFNGLSDMNATSDVNV 1046
```

```
QY 1054 QQLSDTSVLVPIPNWSQVSQQTVPQNYRYLRTARKGVDGYYVLIIRDGNQTTLTLP 1113
DB 1047 QQLSDTSVLVPIPNWSQVSQQTVPQNYRYLRTARKGVDGYYVLIIRDGNQTTLTLP 1106
QY 1114 NICDDDTGVLSDAQTSYITKTFEFTPSTBQWMDMSBTEGVFNIESVELVLEE 1167
DB 1107 NICDDDTGVLSDAQTSYITKTFEFTPSTBQWMDMSBTEGVFNIESVELVLEE 1160
RESULT 3
C8BA BACUK
ID C8BA BACUK STANDARD; PRT; 1169 AA.
AC Q45705;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pesticidal crystal protein cry8Ba (Insecticidal delta-endotoxin
DE CryVIIIa(a)) (Crystalline entomocidal protoxin) (134 kDa crystal
DE protein).
GN CRY8BA OR CRYVIII(A) OR 50C(B).
OS Bacillus thuringiensis (subsp. kumamotoensis).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=132267;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-18746 / PS50C;
RA Michaelis T.B., Fonceirada L., Narva K.E.;
RT "Process for controlling scarab pests with Bacillus thuringiensis
RT isolates";
RL Patent number W09315206, 05-AUG-1993.
CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF INSECTS. ACTIVE ON VARIOUS SCARABID BEETLES.
CC -1- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -1- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus
CC -1- SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: U04365; AAA21118.1; --
CC HSP; P07130; LDLC.
CC InterPro; IPR001178; Endotoxin.
CC InterPro; IPR005638; endotoxin_C.
CC InterPro; IPR005639; endotoxin_N.
CC InterPro; IPR008979; Gal_bind_like.
CC Pfam; PF00555; endotoxin_1.
CC Pfam; PF03944; endotoxin_C; 1.
CC Pfam; PF03945; endotoxin_N; 1.
CC Toxin; Sporulation.
CC KW SEQUENCE 1169 AA; 133543 MW; 22EBFCF5BD69909 CRC64;
QY
QY Query Match 52.2%; Score 3152; DB 1; Length 1169;
QY Best Local Similarity 54.0%; Pred. No. 6.4e-172; Indels 84; Gaps 25;
QY Matches 653; Conservative 179; Mismatches 294;
QY 1 MSPNNQNEYIILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSGEGNPELPGNP 60
DB 1 MSPNNQNEYIILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSGEGNPELPGNP 60
QY 61 EFTFIS-STVQTGIGIVGVLGALGVPFAGQIASFYSPVIGOLWPSSTVSUWEMIKQVE 119
DB 61 EFTFIS-STVQTGIGIVGVLGALGVPFAGQIASFYSPVIGOLWPSSTVSUWEMIKQVE 119
QY 61 EFTFIS-STVQTGIGIVGVLGALGVPFAGQIASFYSPVIGOLWPSSTVSUWEMIKQVE 120
DB 61 EFTFIS-STVQTGIGIVGVLGALGVPFAGQIASFYSPVIGOLWPSSTVSUWEMIKQVE 120
QY 120 DLIDQKITDSVRKTAALAGLQGLGDVYQKSLKNWLENRNDTRARSVVVTVQYIALELDF 179
DB 120 DLIDQKITDSVRKTAALAGLQGLGDVYQKSLKNWLENRNDTRARSVVVTVQYIALELDF 179
```

Db 121 ELINQIAEYARNKALSELGLGNNYQILYLTALBEMKENPNSPALDRVRNRPILDSLF 180
Qy 180 VAKIPSPAIQORVPLSVYAQAANLHLLLRDASIFGAEWGFTPGISTFYDQVTRTA 239
Db 181 TOYMPSPRVNPFVFLTVYVYMAANLHLLLRDASIFGEENGLSTSTINNYNQMKLTA 240
Qy 240 QYSDYCVKMYNTGLDKLGTNAASWKLKXKHOFRREMTLLVLDLVALPNYDRTYPIETTA 299
Db 241 EYSDHCWKVYETGLAKLGSAAQWIDYNQRRMTLLVLDLVALPNYDRTYPIETTA 300
Qy 300 QLTREVTDPVFNKRETSGGRCRWLSNDSISFSEVASVIRSPHLFDILSEIEFYFTRA 359
Db 301 QLTREVTDPVFNKRETSGGRCRWLSNDSISFSEVASVIRSPHLFDILSEIEFYFTRA 356
Qy 360 GLPLNTEYLEYVWGHSHIKYQTNWASSALERNYGT-----ITSNKIKYDLANKDIFQVR 414
Db 361 GLPLNTEYLEYVWGHSHIKYQTNWASSALERNYGT-----ITSNKIKYDLANKDIFQVR 414
Qy 357 S--FTSDRYERYWAGHQISYKHIGTSSTFTQMYGTQNLQSTN----FDTNVDIVKTL 410
Db 358 S--FTSDRYERYWAGHQISYKHIGTSSTFTQMYGTQNLQSTN----FDTNVDIVKTL 410
Qy 415 SLGADLAN-----YAAQVGVYASFTLLDKNTGSGVGGTYSKPHHTMQVCTQNTYI 469
Db 416 SLGADLAN-----YAAQVGVYASFTLLDKNTGSGVGGTYSKPHHTMQVCTQNTYI 469
Qy 411 SNGAVLLDIVPGYTYTFFGCPETEPFPMVQNLNTRKT--LTY-KP-ASKDIIIDRTDSE 466
Db 412 SNGAVLLDIVPGYTYTFFGCPETEPFPMVQNLNTRKT--LTY-KP-ASKDIIIDRTDSE 466
Qy 470 DEIPPE--NEPLSGYSHLSHITSYSFKNASPARYGNLPVPAWTHRSADVTNVTYSD 527
Db 471 DEIPPE--NEPLSGYSHLSHITSYSFKNASPARYGNLPVPAWTHRSADVTNVTYSD 527
Qy 467 LELPETSQPNYESYSHRGLHIT--FYSSSTST-----YVPVSWTHRSADLTNTVWSG 520
Db 468 LELPETSQPNYESYSHRGLHIT--FYSSSTST-----YVPVSWTHRSADLTNTVWSG 520
Qy 528 KITQIPVKAHTLVGTTVIKPGFTGNNILKRTSSGPLATYSVSVKSPLSQRYARIRY 587
Db 529 KITQIPVKAHTLVGTTVIKPGFTGNNILKRTSSGPLATYSVSVKSPLSQRYARIRY 587
Qy 521 BITQIPGKSTIGENTVYIKRGVGTGDLVALTDR--IGSCBQFMIPFESQRFIRY 578
Db 522 BITQIPGKSTIGENTVYIKRGVGTGDLVALTDR--IGSCBQFMIPFESQRFIRY 578
Qy 588 AS--TTNRLP--VYISGTRIYSINVKT--MKKGD-DLTFNTPDLATGTATFNSYDSL 642
Db 589 AS--TTNRLP--VYISGTRIYSINVKT--MKKGD-DLTFNTPDLATGTATFNSYDSL 642
Qy 579 ASNETSVISGLNQSGT-----LKFQNTYSKNENDLTYND-----FKYIYPRVI 625
Db 580 ASNETSVISGLNQSGT-----LKFQNTYSKNENDLTYND-----FKYIYPRVI 625
Qy 643 TVGADS-----PASGEVYV--DKPELIPVNATERAEEDLDVAKAVNGLFTSKDAL 693
Db 644 TVGADS-----PASGEVYV--DKPELIPVNATERAEEDLDVAKAVNGLFTSKDAL 693
Qy 626 SVNASSIQRISIGIQNTNLFILDRIFIPVETYEATLEDEAKAVNGLFTSKDAL 685
Db 627 SVNASSIQRISIGIQNTNLFILDRIFIPVETYEATLEDEAKAVNGLFTSKDAL 685
Qy 694 QTSYTDYQVNOANLVECLSDLYPNKRMWDAVKEARLVQARNLLQDTGFNRINGEN 753
Db 695 QTSYTDYQVNOANLVECLSDLYPNKRMWDAVKEARLVQARNLLQDTGFNRINGEN 753
Qy 686 QPGVTDYEVNOANLVECLSDLYPNKRMWDAVKEARLVQARNLLQDTGFNRINGEN 745
Db 687 QPGVTDYEVNOANLVECLSDLYPNKRMWDAVKEARLVQARNLLQDTGFNRINGEN 745
Qy 754 GWTSGTGLEAVGDVLFKORSRLTSAREIDTETPTLYLOQIDSLIKPTRYKLGPI 813
Db 755 GWTSGTGLEAVGDVLFKORSRLTSAREIDTETPTLYLOQIDSLIKPTRYKLGPI 813
Qy 746 GWTASTGIEVTEGDAVFGKRYLFGAREIDTETPTLYLOQIDSLIKPTRYKLGPI 805
Db 747 GWTASTGIEVTEGDAVFGKRYLFGAREIDTETPTLYLOQIDSLIKPTRYKLGPI 805
Qy 814 GSSODLEIKLIRHANOIVKVPNNLLPDVLPVNSCGGIDRCSCQVYDAMLALNNGEN 873
Db 815 GSSODLEIKLIRHANOIVKVPNNLLPDVLPVNSCGGIDRCSCQVYDAMLALNNGEN 873
Qy 806 GSSGLEIYTHQTNRLVKNVPPDOLLDPVFPVNDGRINRCSEKQVNSRLEVENR--- 862
Db 807 GSSGLEIYTHQTNRLVKNVPPDOLLDPVFPVNDGRINRCSEKQVNSRLEVENR--- 862
Qy 874 GNMSSDSHAFSHIDTGBIDNENTGIWVVKIPIPTNGYATLGNLELVVEGPLSGETLER 933
Db 875 GNMSSDSHAFSHIDTGBIDNENTGIWVVKIPIPTNGYATLGNLELVVEGPLSGETLER 933
Qy 863 ---SGBAEHFSIPIDTGBELDNAGIWWGPKITDPEGYATLGNLELVVEGPLSGDALER 919
Db 864 ---SGBAEHFSIPIDTGBELDNAGIWWGPKITDPEGYATLGNLELVVEGPLSGDALER 919
Qy 934 AQOQEQWODVAKRGASEKAYAAQAIIDRLFADYQDQKLSNGVEMSDMLAQLVQS 993
Db 935 AQOQEQWODVAKRGASEKAYAAQAIIDRLFADYQDQKLSNGVEMSDMLAQLVQS 993
Qy 920 LQKEEQWQKIQWTRERRETDRYMAKQAVDLYADYQDQKLSNGVEMSDMLAQLVQS 979
Db 921 LQKEEQWQKIQWTRERRETDRYMAKQAVDLYADYQDQKLSNGVEMSDMLAQLVQS 979
Qy 994 IPYVNDALPEIICGMYNTSFTELTNLOQANMLYDLRNAIENGDFRNLGSDNNTSDVNV 1053
Db 995 IPYVNDALPEIICGMYNTSFTELTNLOQANMLYDLRNAIENGDFRNLGSDNNTSDVNV 1053
Qy 980 IPYVNEFPPIPGNYTKFTLDRLOQANGLVDORNAIENGDFRNLGSDNNTSDVNV 1039
Db 981 IPYVNEFPPIPGNYTKFTLDRLOQANGLVDORNAIENGDFRNLGSDNNTSDVNV 1039
Qy 1054 QQLSDTSVLVLPNNNSQVSOQFTQPNRYVYVLRVTRAKBGVGDGVYIRDGANQETLTF 1113
Db 1055 QQLSDTSVLVLPNNNSQVSOQFTQPNRYVYVLRVTRAKBGVGDGVYIRDGANQETLTF 1113
Qy 1040 QQINHTSVLVLPNNNEQVSKFTQPNQRYVYVLRVTRAKBGVGVGVYIRDCGQSETLTF 1099
Db 1041 QQINHTSVLVLPNNNEQVSKFTQPNQRYVYVLRVTRAKBGVGVGVYIRDCGQSETLTF 1099
Qy 1114 NICDDTGVLSADQTS-----YIT-----KITVETPSTEQWIDHSETGEPNI 1157
Db 1115 NICDDTGVLSADQTS-----YIT-----KITVETPSTEQWIDHSETGEPNI 1157
Qy 1100 SASDYDNGMYDTQASNTNGNTNTSVYMKPAISKRTVDISSVYQNMWIEISSETGYFI 1159
Db 1101 SASDYDNGMYDTQASNTNGNTNTSVYMKPAISKRTVDISSVYQNMWIEISSETGYFI 1159
Qy 1158 ESVELVLREE 1167
Db 1159 ESVELVLREE 1167
Qy 1160 ESVELVLDVE 1169
Db 1161 ESVELVLDVE 1169

RESULT 4

C9CA_BACTO STANDARD; PRT; 1157 AA.
ID C9CA_BACTO STANDARD; PRT; 1157 AA.
AC 045733;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DB Pesticidal crystal protein cryCa (insecticidal delta-endotoxin
CryIXC(a)) (Crystalline entomocidal protoxin) (130 kDa crystal
protein).
GN CRY9CA OR CRYIXC(A).
OS Bacillus thuringiensis (subsp. tolworthi).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1442;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=BT02518A; PubMed=8572715;
RX MEDLINE=96141404; PubMed=8572715;
RA Lambert B., Buyese L., Decock C., Jansens S., Pien S., Saey B.,
Seuring J., Van Audenhove K., Van Rie J., Van Vliet A., Peteren M.;
"A Bacillus thuringiensis insecticidal crystal protein with a high
activity against members of the family Noctuidae.";
RL Appl. Environ. Microbiol. 62:80-86(1996).
CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
EPITHELIAL CELLS OF LEPIDOPTERA LARVAE. HAS A FAIRLY BROAD
SPECTRUM OF ACTIVITY AGAINST MEMBERS OF THE FAMILIES PYRALIDAE,
PLUTELLIDAE, SPHINGIDAE, AND NOCTUIDAE. IT IS THE FIRST
INSECTICIDAL CRYSTAL PROTEIN WITH ACTIVITY AGAINST CUTWORMS. NO
ACTIVITY IS OBSERVED AGAINST SOME BEETLES, SUCH AS COLORADO POTATO
BEETLE.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
sporulation and is accumulated both as an inclusion and as part of
the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; Z37527; CAA85764.1; -.
CC PIR; A59350; S49247.
CC HSP; P07130; IDIC.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin C.
DR InterPro; IPR005639; endotoxin N.
DR InterPro; IPR008979; Gal_bind_Like.
DR Pfam; PF00555; endotoxin_1.
DR Pfam; PF03944; endotoxin_C_1.
DR Pfam; PF03945; endotoxin_N_1.
KW Toxin; Sporulation.
SQ SEQUENCE 1157 AA; 129775 MW; C364391EF7FDFB8A CRC64;

Query Match 51.7%; Score 3123; DB 1; Length 1157;
Best Local Similarity 52.3%; Pred. No. 2.8e-170;
Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;
Qy 1 MSPNNQNEYELDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSGEENPELFGNP 60
Db 1 MNRNNQNEYELIDAPHCCGSPDDDDVRYPLASDPNAAALQNNYKDYLTQDDEYDYNP 60
Qy 61 ETPFSS-STVGTGIGVQVGLGALGVPPAGQIASFYSFVQGLWPSSTVSVWEMTKQVE 119
Db 61 SLSISGRDAVQTALTWGRILGALGVPPSGQIVSPYQFLNTLTWPNVDTAIWEAFMQVE 120
Qy 120 DLIDQKTDTSVRKTAALAGLQGLDGLDVTYQKSLKNWLENRNDTRARSVVVQYIAELDP 179

```
Db 121 ELVNOQITEFARNQALRLQGLGDSFNVYQSRSLQNLADENDTRNLVSVRAQPIALDLDF 180
Qy 180 VAKIPSPALSGQSVPLLSVYQAANLHLLLRDASIFGAEWGTPGKEISTEYDRQVTRTA 239
Db 181 VNAIPFPAVQGVPLSVYQAANLHLLLRDASIFGAEWGTPGKEISTEYDRQVTRTA 240
Qy 240 QYSDYCVKRYNTGLDKGTNAASWLKTHQPREMTLLVLDLVALPNNYDTRTYPIETTA 299
Db 241 KTYNCETWYNTGLDLRGNTESWLRYHQPREMTLLVLDLVALPNNYDTRTYPIETTA 300
Qy 300 QLTREVYTDPIVFNRTSGGFCRRWSLNSDISFSEVESAVIRSPHLPDILSEIEPTTTRA 359
Db 301 QLTREVYTDPIVFNPPANVGLCRRWGTNFPYNTSELENAPFIRPHLPDLRLNSLTISNR- 359
Qy 360 GLPLANNTEYLEVWGHVSIKYKNTNASSALERNYGTITSNKIKY---YDLANKDIFVRSRL 416
Db 360 -FPV-SSNPDYWGHTLRSYLNDSAVQEDSGLTITRATNPGVDGTR-----TEST 413
Qy 417 GADLANYYAQVYGPVASFPL--LDKNTGSGVGGFTYSKPHHTTMQVCTQNTYNTIDBEP 474
Db 414 AVDPFRSALIGYGNRASFPVGLFNGTTPANGG-----CRDLYTNDLPLP 461
Qy 475 ENEPLSRGYS-HRLSHITSFSPKN-ASSPARYGNLPVPAWTHRSADVTYTVYSDKLTQI 532
Db 462 DE---STGSSTHRLSHVTPFPFQTNQAGSIANAGSVPTVYVTRRDVLDNNTITENRITOL 518
Qy 533 PVVKANTLVSGTIVKPGTGNILKRTSGGLAVTSVSKPLSGVPLSORVBARIRYASTTN 592
Db 519 PLVKAAPVSGTIVKPGTGGILKRTNGTNGFGLRVTVNSPLTQOYKLRVFPASTGN 578
Qy 593 LRLPVTISGTRIYSINVMKMGDDLTENTF---DLATIG---TAPTFSNYSDSLTVGA 646
Db 579 FSIKVLGGVSGIDVRLGSTMNGQBLTYESFPTRBTPTTTPGPNPPTPTQAQELTVNA 638
Qy 647 DSPASGEVYVDFKPLIPVNAFEAREBDLVAKAVNGLFTSKKDALQTSVTVYQVQAA 706
Db 639 EGVSTGGEYVIDRIEIVPNPAREAREDELEAKKAVASLFTFRTRDGLQVNVTVQVQAA 698
Qy 707 NLVECLSDLEYPNKEMLDVAKKRLVQARNLQDTPGNRING--ENGWGTSGTGLEVA 764
Db 699 NLVSCLSDEQGHQKMLLBAVRAAKLSRERILLQDPDFTNTSTEENGWKAANGVTIS 758
Qy 765 EGVLFKDRSLRLTSAREIDTETPTLYLQOIDEKLKPYRYLKLKPGFSGSQDLKILKI 824
Db 759 EGGPFKGRALQASAR---ENYPTVYQKVDASVLKPYRYRLDGLGVKSSQDLKILKI 814
Qy 825 RHANQIVKVPNNLLPDVLPVNSCGGIDRCSEQQYVDANLALNENGENGMSDSEHAPS 884
Db 815 HHKHLVHLVKNVPLNSDYSYSDGSCSGINRCDEQHQVDMQDLABHHHPMDCBAAQTBEFS 874
Qy 885 FHIDTGEIDLNENTGIWVVPKIPPTNGYATLGNLELVEEGPLSGETLERAQQOQWODK 944
Db 875 SYINTGDLNASVQGIWVVLKVRITTDGATLGNLELVEEGPLSGESLREQRONAKWNAE 934
Qy 945 MARKRGASKEYAQAQIDRLFADYQOQKLNSEVMSDMLAQNLYQSIPTVYNDALPE 1004
Db 935 LGRKRAEIDRVYLAQAQINELFDYQOQKLNSEVMSDMLAQNLYQSIPTVYNDALPE 994
Qy 1005 IPGANTYFTLNRQAQANLIDLNAIPNGDFRNLGSOWNATSDVNVQOQSDTSVLVI 1064
Db 995 IPGINBYITELSDRLQAQSYLTSRNAVQNGDFNSGLDSWNTTMDASVQDQGNHFLVL 1054
Qy 1065 PNWNSVQSQFTQVPPYRYVLTARKEGVDGVIIRDGANQVETLTFTNICDDDTGVLS 1124
Db 1055 SHWDAQVQSQLRVNPENCKYLRVARKVGQDGYVITRDGAHQHETLTFTNACDYVNGTY 1114
Qy 1125 ADQTSYTKTVEPTFTEQVHIDMSSETGVFNIESVELVLEE 1167
Db 1115 VNDNSYITEEVFPETKHMVVESESGSFYDSIEFIETQE 1157
```

RESULT 5

```
C9EA_BACTA STANDARD; PRT; 1150 AA.
AC 09ZNL9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pesticidal crystal protein cry3a (insecticidal delta-endotoxin
DE CRYIXE(a)) (crystalline entomocidal protoxin) (130 kDa crystal
DE protein).
GN CRY9EA OR CRYIXE(A).
OS Bacillus thuringiensis (subsp. aizawai).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1433;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SSK-10;
RA Midoh N., Oyama K.;
RT "Bacillus thuringiensis cry gene for insecticidal crystal protein.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF INSECTS.
CC -I- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -I- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -I- SIMILARITY: Belongs to the delta endotoxin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB011496; BAA34908.1; -.
DR HSSP; P07130; IDLC.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal_bind_Like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
KW Toxin; Sporulation.
SQ SEQUENCE 1150 AA; 129895 MW; 7D6AB93D6BDC97EB CRC64;

Query Match 46.7%; Score 2825.5; DB 1; Length 1150;
Best Local Similarity 48.6%; Pred. No. 2.6e-153;
Matches 574; Conservative 197; Mismatches 364; Indels 47; Gaps 17;

Qy 1 MSPNNQNEYILDASSSTSVSDNSVRYPLANDQPTTLQNNMYKDYLRMSGE-----NP 54
Db 1 MNRNPNNEYIIDAPYCGSCFSDDDVRYPLASDPNRAQNNMYKDYLYQDYDGTGSLNP 60
Qy 55 ELFGNPEPTFISSTVTQGTIGVQVLGALGVPFAGQIASFYSGVQLWPSSTVSWEM 114
Db 61 NLSINPR-----DVLQGTINIVGRILGLFSLGVPFAGQLVTFYTLNQLNLPNDNAWEAF 115
Qy 115 MKQVEDLIDQITDTSVTKTALAGLQGLDGLDVTQKSLKNLNRNDTRARSVVVQYIA 174
Db 116 MAQTEELIDQKISAQVVRNALDGLTGLDYEEYEALEELERPNAGARA-NLVQRFEN 174
Qy 175 LELDFVAKIISPALS-CQE---VPLLSVYQAANLHLLLRDASIFGAEWGTPGKEISTP 230
Db 175 LHTAFVTRMPSFGPGSGORDAVALLTVYQAANLHLLLRDASIFGAEWGTPGKEISTP 234
Qy 231 YDRQVTRTAQYSDYCVKRYNTGLDKGTNAASWLKTHQPREMTLLVLDLVALPNNYD 290
Db 235 FNAQOEIRIYTHCHVETVYARGLEDVRGTWNTSENLNHYRFRREMTLMAQDLVALPNNY 294
Qy 291 RTYPIETTAQLTRVYTDPIVFNRTSGGFCRRWSLNSDISFSEVESAVIRSPHLPDILS 350
```

Db 295 ROYFNGANPOLREIYTDIVNPPANQICRWGNPNYNTSELENAPRPHPLERLN 354
Qy 351 EIEFTYTRAGLPLANNTEYLYVWGHISIKYKNTNASSALERNYGTITSNKKYKYDLANKDI 410
Db 355 RLTISSRNRYTAPTNS-FLDYSNGHTLQSHANNPTVYTSYGQITSN-TRLPNTTN-CA 411
Qy 411 FOVRISGLADLANVAQVGPVYASFTLLDKNTGSGVGFTYSKHTTMOVCYNTVID 470
Db 412 RAIDSRARRNGLYANLYGV--SSUNIPP--TG-----VMSRITNAANTCRQDLITTE 460
Qy 471 EIPPENEPLSRGSHRLSHITSYSPSKNASP-ARYGNLPVPAWTHRSADVNTVYSOKI 529
Db 461 ELPLENNP-----NLLSHVTLRFTNTQGGPLATLGPVTVVTRVEDVFTTITADRI 515
Qy 530 TQIPVKAHTLVGTTVINGPFTGNILKRTSSGLPLATSVSVKSPISQRYRARIAS 589
Db 516 TOLPVPVYKASEIGGTTVVKPGFTGDLRRTDGGAVGITIRANVAPLTQOYRIRLYAS 575
Qy 590 TTN--LRLPVTISGTRIYSINVKMTWKGDDLTENTDLATIGTAPTESNYSDSLTVGAD 647
Db 576 TTSFVNLPVNSAA---GFTLPSTMAQNSLTIESFNTLEVTHTIRFSQSDTTLRLMIP 632
Qy 648 SPASGGEVYVDPKFLIPVNAFEABEDLVAKKAVNELFTSKDALQTSVTDYQVNOAAN 707
Db 633 PSISQGEVYVYDKLIPVNPTRAEAEDELDAKKAVASLFTTRDGLGVNNTDYQVNOAAN 692
Qy 708 LVECLSDLELYPNEKMLMDAVKAEALVQARNLQOTGPNRING--ENGWGTSGTGIEVAB 765
Db 693 LVSCLSDEQYGHDKOMLAEVAAKRLSRERLLQDPDFNEINSTENGWKAASNGVTISE 752
Qy 766 GDVLPKDRSLRATSAREIDTETPTVLYQOIDESELLKPVTRKLAGPIGSSQDLKILIR 825
Db 753 GGPFFKGRALQASAR----ENPTTYIQKVDASTLKPTRYKLDGFPVQSSODLSIDLII 808
Qy 826 HRANQIVKRNPNLLFDVLPVNSCGGIDRCSEQQYVDANLALENNGENGNSSDSHAFSP 885
Db 809 HKHVHLVKNVDPNLVSDYSDGSCGINRCBEOHQVDVQLDAEDHPKOCCEAAQTHEFSS 868
Qy 886 HIDTGEIDLNENGTWVFKIPPTNGYATLGMLELVEGGPLSGFTLEKRAQOQBOQDKM 945
Db 869 YHTGDLASVDQGIWVYQVTRTDGATLGNLELVEVGPLSGESLREQRDNKXNBEV 928
Qy 946 ARKRGASKAYYAAQADIRLFPADYQDKLNSGVMSDMLAAQNLVQSTIPVYNDALPRI 1005
Db 929 GKRAETDRIQDAQAQNLHLPVDYQDQQLSEVGHADIIDQNLIASISDVISDAVLOI 988
Qy 1006 FGMNYSPTFLNRLQAAWLYDLRNLAIPIKNGDPRNGLSDMNATSDVNYQQLSDTSVLVIP 1065
Db 989 PGINEMYTELSNRLQQAASLYTSRNVYVQNGDFNSGLDSWNATDTAVQDGNMHEFLVS 1048
Qy 1066 NWSQVSOQTPVQPNRYRVLRTAKRKGVDGYVIRIGANQOTELTFTNI CDDDTGVLSA 1125
Db 1049 HMDAQVSOQFRVQPNCKVLRVTAKKVNGDGYVTIQDGAHRETLTTFNACDYDNGTHV 1108
Qy 1126 DQTSYITKVTPTSTOVIDMSSTEGVFNLSKVELVLEE 1167
Db 1109 NDNSTITKELVPYPTHEHMEVSETEGTYTIDSTIEPTQE 1150

RESULT 6

C9DA_BACTP
ID C9DA_BACTP STANDARD; PRT: 1169 AA.
AC 006014;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Peatidial crystal protein cry9Da (insecticidal delta-endotoxin
DE CryIXD(a)) (crystalline entomocidal protoxin) (132 kDa crystal
protein).
GN CRY9DA OR CRYIXD(A).
OS Bacillus thuringiensis (subsp. japonensis).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

NCBI_TaxID=128936;
OX [1]
RN SEQUENCE FROM N.A.
RC STRAIN=N141;
RA Asano S.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF INSECTS.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D85560; BAA19948.1; -.
DR HSP; P07130; 1DLC.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin C.
DR InterPro; IPR005639; endotoxin N.
DR InterPro; IPR008979; Gal bind_like.
DR Pfam; PF00555; endotoxin_1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
DR Toxin; Sporulation.
DR Toxin; Sporulation.
SQ SEQUENCE 1169 AA; 132228 MW; 659AB257229DE5E9 CRC64;
Query Match 44.9%; Score 2715; DB 1; Length 1169;
Best Local Similarity 46.9%; Pred. No. 5.3e-147;
Matches 558; Conservative 204; Mismatches 383; Indels 46; Gaps 17;
Qy 1 MSPNNQNYEILDASSSTVSNSVRYPLANDQTTTLQNNYKDYLRMSEGENPELFGNP 60
Db 1 MNRNNQNEYVIDAPHCPCPADVVYKPLTDDPNAGLQNNYKEYLQTYGGDYTDPLNP 60
Qy 61 EFTFIS-SSTVGTGIGVQVGLGALVPPAGQIASPVSPVQGLPSSSTVSVMEMIKOYE 119
Db 61 NLSVSGKDVQVGINIVGLRLSPFPFPSSQWVTVYLLNSLNFDDNSVMDAPRVE 120
Qy 120 DLIQKRTDSVRKTALAGLQGLDGLDVYQSKLKNLNRNDTRARSVVVYQYIALEDF 179
Db 121 ELIDQKISEAVKGRALDDLTLGLQVNYLVVEALDEMLNRPNGARA-SLVSRFNILDSLF 179
Qy 180 VAKTSPALSG-----QEVPLLSVYAOAANLHLLILRLDASIECAEWGTFPGHISTPYDRQ 234
Db 180 TQFMPSPG-SGPGSQNYATILLPVYAOAANLHLLILKADIIYGARWGLNQTDIQPHSRQ 238
Qy 235 VTRTAQYSDYCVKVTNTGLDKLKNAAASWLYHOFREMLLLDLVALFPNYDRTYTP 294
Db 239 QSLQTYTNHCVTAINDGLAELRGTAESNFKYQYRREMLTAMDVLVALFPYNLQYTP 298
Qy 295 TETTAQLTRVYTDPIFVN--RETSGGFCRWLSN-----SDTSFSEVSASVIRSHPLFDI 348
Db 299 DGTNPQLTRVYTDPIAFDPLBQPTTQLCRSWYINPAFRNHLNFSVLENSLRIRPPLPER 358
Qy 349 LSEIEP----YTRAGLPLNTEYLEYVWGHISIKYKNTNASSALERNYGTITSNKKYDYD 405
Db 359 LSNLQILVNYQTNGSA-----WRGSRVRYHLYHSSIIQEKSYGLLSDPVGANTKNV 408
Qy 406 ANKDIQVRSGLADLANVAQVGPVYASFTLLDKNTGSGVGFTYSKHTTMOVCYNTQ 465
Db 409 QNNDIYQIIISOVSNPASPVGSSYSWDTNRYL-----SSGQVSGISGTYQOQIPACVLOQ 463
Qy 466 YNTIDEIP---PENEPLSRGSHRLSHITSYSPSKNAS-SPARY-GNLPVPAWTHRSADV 520
Db 466 YNTIDEIP---PENEPLSRGSHRLSHITSYSPSKNAS-SPARY-GNLPVPAWTHRSADV 520


```
Db 464 RNSTBELPSLNPEDG-IIRNYSHRLSHITQYRFQATQSGSPSTVSANLFTCVWTHRDVUL 522
Qy 521 TMTVSDKITQIPVKAHALVSGTIVIKGPGFTGGENILKRTSGPLAYTSVSKSPLSQR 580
Db 523 DMTITANQITQPLVKAELSSGATVVGKPGFTGGDVIRRTTGGGAIKRVSVTGLTQR 582
Qy 581 YRARIRYASTTNLRPLFTVITSGTRIYSINNVKNKNGDDLTFTNFDLATIGTATFPNYSQ 640
Db 583 YAIRPRYASTIDFPFVTRGGTITNNFRTRTNWQGESRYESRVTEPTTPEFTQSQD 642
Qy 641 SLTVGADSPASGGEVTVDFKELIPVNAPEAEEDLVAKKAV-NGIPTSKEQALQTSVD 699
Db 643 IIRTSQISGNGEVLEIDRIEIIIPVNPAPAEEDLEAKKAAKQNLFTTRDGLQNVVD 702
Qy 700 YOVQANALVBCLSDELYPEKRMKMDVAKKLVQARNLQDTGFNRING--ENGWTG 757
Db 703 YQVQDQANLVCLSDROYGDKKLEAVRAAKELSRNLLQDPDPTINSTENGWKA 762
Qy 758 STGIEVAGDVLFPKDRSLRLTSAREIDTETPTLYLQYQIDESLLKPYTRYKLGKPTGSSQ 817
Db 763 SNGVTISEGPPFKGRALQASAR----ENYPTIYQKVDASVLKPYTRYLDGFKVSSQ 818
Qy 818 DLEIKLIRANQIVKGVNDLLPDLVLPVNSCGIDRCSEQQYVDANLANNGB-NGNM 876
Db 819 DLEIDLIHYKHLVKNVNDLVSDTSDGSCGMNRCBEQQMVNAOLETSEHHHPMDCE 878
Qy 877 SSDSHAFSPHIDTGEIDLNTGICWVFKIPTNGYATLGNLELVEGLSGELTERRAQ 936
Db 879 AACTHFFSYINTGDLNASVDQGIWVVKVTRTDGTATLGNLELVEGLSGELTERRAQ 938
Qy 937 QBOQWQDMARKGASAKAYAAKQAIIDLFPADYQDOKLNSGVEMSDMLAAQNLVOSIPY 996
Db 939 DNAKNNAELGRKRAEDRVYLAQKQINHLFVDYQDQQLNPETGLAEINAEASLVESISG 998
Qy 997 VYNDALPEIPGMYTSTFELTRLOQANLYDLRNALPNCDEPGLSDNATSDVNVQOL 1056
Db 999 VYSDTLQIQPGINYEIYITELSDLOQASYLYTSRNAVQNGDFNSGWNITTDASVQOD 1058
Qy 1057 SDTSVLVIPNNSQVQQFTQVQPNRYALRVLTARKEGVGEGYVIRIDGAMQETELTFNIC 1116
Db 1059 GNWHLVLSHWDQVSOQLRVNPNCKYLVLRVTKRVGGGEGYVIRIDGAHQETLTFNAC 1118
Qy 1117 DDDTGVLSADQTSYITKTVEPTSTPQVWIDMSSETGQVNVIESVRLVLRBEE 1167
Db 1119 DYDVNGTYVNDNSYITEEVVFYPTKHMVVESESGSYDSIRPIETQE 1169

RESULT 7
C1KA_BACTM STANDARD; PRT; 1215 AA.
ID C1KA_BACTM AC Q45715;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Pesticidal crystal protein cryiKa (insecticidal delta-endotoxin
DE CryiK(a)) [Crystalline entomocidal protoxin] (137 kDa crystal protein).
GN CRYIKA OR CRYIK(A) OR CRYIK.
OS Bacillus thuringiensis (subsp. morrisoni).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1441;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F190;
RX MEDLINE=96102856; PubMed=8586263;
RA Koo B.T., Park S.-H., Choi S.-K., Shin B.S., Kim J.I., Yu J.H.;
RT "Cloning of a novel crystal protein gene cryiK from Bacillus
RT thuringiensis subsp. morrisoni.";
RL FEMS Microbiol. Lett. 134:159-164(1995).
CC -!- FUNCTION: PROMOTES COLLOIDOSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF INSECTS. SELECTIVELY TOXIC TO ARTOGEIA RAPAE
CC AND NOT ACTIVE ON PLUTELLA XYLOSTELLA.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
```

```
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U28801; AAB00376.1; -.
CC HSSP; P02965; 1CIY.
CC InterPro; IPR001178; Endotoxin.
CC InterPro; IPR005638; endotoxin_C.
CC InterPro; IPR005639; endotoxin_N.
CC InterPro; IPR008979; Gal_bind_like.
CC Pfam; PF00555; endotoxin; 1.
CC Pfam; PF03944; endotoxin_C; 1.
CC Pfam; PF03945; endotoxin_N; 1.
CC Toxin; Sporulation.
CC SEQUENCE 1215 AA; 137378 MW; 7CAP7A3311893D9B CRC64;
Query Match 44.9%; Score 2714; DB 1; Length 1215;
Best Local Similarity 46.3%; Pred. No. 6.4e-147;
Matches 581; Conservative 194; Mismatches 351; Indels 130; Gaps 23;
Qy 1 MSPNNQNEVEILDASSSTVSQNSVRYPLANDQTTTLQNNYKYDLMSGENPCLPCNP 60
Db 1 MNSNRKNEHEITNALSIAPVNSHQMLSPD-----ARIDSLCVAEG-----NNI 47
Qy 61 ETPISSTVQTGIGVGVGLGVPFAGIAPFISFIVGQALPSSVTVVEMIMKQVED 120
Db 48 DPFVSAVSTVQTGISIAGRILGVLPFAGQALASFYSLVGLWELPFSGR-DPWEIPEWVEQ 106
Qy 121 LI-DQKITPSVRKTALAGLQGLGDLVYQKSLKNMLENDRTRARVVVTVYIALELDP 179
Db 107 IVRQQTITDSVRDTARLEGLGRVRSYQQALETYLDNENDARSRIIRERYIALELDI 166
Qy 180 VAKIPSPALSGQEVPLISVYQAANLHLALLRLDASIFGAEMGPTPGISTFYDQVTRTA 239
Db 167 TTAIPFISIRNBEVPLLVYQAANLHLALLRLDASLFGSEMGSSADVQYQIQRYTE 226
Qy 240 QISDYCVKYNVTGLDKLGTNAASWLKTHQFREMTELLVLDLVALFPNYDTRTYPIETTA 299
Db 227 EYSNHCQWYNTGLNRLRGTTAETWRYNQFRDLTLGLVLDLVALFPNYDTRTYPIETTA 286
Qy 300 QLTREYVYTD--IVFNRETSGGFCRRWLSNDSISFSEVESAVITRSPHLDLSIEFTTT 357
Db 287 QLTREYVYTDPNGVWAGPNNS-----WFRNG-ASFSAIENAIIRQPHLYDPLTNLTITR 339
Qy 358 RAGLPLANTYLEYFWYVGHISIKYENTWASSALERNYGTITSNKIKYYDL--ANKDIFQVRS 415
Db 340 RSQV---GTTIMNLWAGHRIITFNRIQGGSTSEMYGAIT-NPVSVDIPFNVNDIVRTVS 395
Qy 416 LGADLANYVAQVYVYASFTLLDKNTKTSVSGVGGFTYSKPHITTMQVCTQNTYNTIDEIPPE 475
Db 396 LAGGLGSLSGIRYGLTRVEDFDMIFRN-HPDIVTGLFVHPGHA--GIATQVKDSDELPEPE 452
Qy 476 --NEPLSRGYSRHLSHITVSFSPKNSASSPARYGNLVPFANTHRASADVNTVYSKLTQIP 533
Db 453 TTEQPNYRAPFSLHLSHI-----SMGPTTQDVPFVYVSWTHOSADRTNTINSRITQIP 504
Qy 534 VYKAHTLVSGTTVIKGPGFTGGMILKRTSGGLAYTSVSKSPLSQRYARIRVASTTNL 593
Db 505 LVKAHTLQSGTTVVKGPGFTGGDILKRTSGGPFAPFNVDLFLNLSQRYARIRVASTTNL 564
Qy 594 RLFPVITSGTRIYSINNVKNKNGDDLTFTNFDLATIGTATFPNYSLSLTVGADSPASGG 653
Db 565 RIYTVAGERIIPAGQFDKTMADAGPLTFQGSFVSATINTATFPERSSSLTIGADTFSSGN 624
```



```

Qy 654 EYVDKFEIIPVATFEAEEDLVAKKAVNGLFSTKDD-ALQTSVTDYQVNNQANLVECL 712
Db 625 EYVDKFEIIPVATFEAEEDLVAKKAVNGLFSTKDD-ALQTSVTDYQVNNQANLVECL 684
Qy 713 SDELYPNEKRLMDAVKAEKLVQARNLQDTGKRLNG--ENGWTSSTGIEVAGDVL 770
Db 685 SDEFLCKRRELEVKYAKRLDERNLQDTFTSISGQDRGWSIGSIOGGDDIF 744
Qy 771 KDRSLRLSAREIDETETPTLYQOQIDESLKKPYTRYKLGKFGISSDLEKLIRHRANQ 830
Db 745 KENVRLPGT--VD-ECYPTLYQKIDESQLKSYTRYQRCYVIEDSQDLLEILYRYNAKH 801
Qy 831 IVKRVPNMLP-----DVLVNSCGIDRCSCQQVVDANLALNNGENGMS-SSSHAFSP 885
Db 802 ETLSPVPGTESPWPSSGVVPSRCGEPNRCAPR--IEMNPDLDCCRYGKCKVHSHFSL 859
Qy 886 HIDTGEIDLNENTGIWVFKLPTTNGVATLNLALVEBGLSGETLERAOQOQWODKM 945
Db 860 DIDVGTDLNEDLGWVIFPKLQDGHAKGLNLEFIEKPLGLKALSRVRAEKWRDKY 919
Qy 946 ARKRGASEKAYAAKQAIIDELFADYQDQKLSGVEMSDMLAAQNLVOSIPYVYNALPBI 1005
Db 920 EKLQETKRVYBAKESVDALFVDSQYDKLQANTNIGIIGHGADQVHREIPYSELPLVI 979
Qy 1006 PGMVYTSFTELTNRLQOANLYDLRNLPAIENGDFRNLGSDNATSDVNVQQLSDTSVLVIP 1065
Db 980 PSINAAIFEELEGGHIFKAYSILYDARNVIKNGDFNGLSCNNVKGHDVQVQNHHSVLVS 1039
Qy 1066 NNNQSVQOQFTVQPNRYVVLRTARKEGVDGYYLIRDGANOTETLTFN----- 1114
Db 1040 EWEAEVSQKRVCPDRGIVLRTAYKGGYGGCVTIFHEFEDNTDLKFRNVEHEVPNN 1099
Qy 1115 --ICDDDTGVLSADQTS----- 1129
Db 1100 TVTCNDYTNQSAEGSTDACNSYNGVEDYENRYEPNAPSAPVNYTPYRGMVTDQGY 1159
Qy 1130 -----YIKTVETPTSTQWIDMSHETGVPNIESVELVLBBE 1167
Db 1160 NHCVSDRGYRNHTPLPAGYVTLLELYETPETEQVWIEIGETGTFTVGSVELLMBE 1215

RESULT 8
ID C1BB_BACTU STANDARD; PRT: 1229 AA.
AC Q45739;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Pesticidal crystal protein cryIbB (insecticidal delta-endotoxin
GN CryIbB) (Crystalline entomocidal protoxin) (140 kDa crystal protein).
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-21110 / EG5847;
RA Donovan W.P., Tan Y., Jany C.S., Gonzalez J.M. Jr.;
RT "Bacillus thuringiensis cryet4 and cryet5 toxin genes and proteins
RT toxic to lepidopteran insects.";
RL Patent number US5322687, 21-JUN-1994.
CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF MANY LEPIDOPTERAN LARVAE.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L32020; AAA22344.1; -.
CC HSP; P02965; ICYI.
CC InterPro; IPR001178; Endotoxin.
CC InterPro; IPR005638; endotoxin C.
CC InterPro; IPR005639; endotoxin N.
CC InterPro; IPR008979; Gal_bind_Like.
CC Pfam; PF00555; endotoxin; 1.
CC Pfam; PF03944; endotoxin C; 1.
CC Pfam; PF03945; endotoxin N; 1.
CC Toxin; Sporulation.
CC SEQUENCE 1229 AA; 139769 MW; A4C949DB675C3269 CRC64;
Query Match 42.2%; Score 2552; DB 1; Length 1229;
Best Local Similarity 43.2%; Pred. No. 1.1e-137; Indels 140; Gaps 22;
Matches 548; Conservative 205; Mismatches 375;
Qy 1 MSPNNQREYELDASSSTSVSDNSVRYPLANDQTTTLQNNYKYDYLMS EGENPELFGNP 60
Db 1 MTSNRKNEEIIALS IPTVSNRSTQNNLSPD-----ARIEDSLCVAE-----VNNI 47
Qy 61 ETPISSTVQVGIGVQVIGALGVPPAGQIATSPYSPVIGQLWPSSTVSVMEMKQVED 120
Db 48 DPPVYASTVQVGINAGRIILGVLPAGQIATSPYSPVIGQLWPSSTVSVMEMKQVED 106
Qy 121 LIDOKITDVRKTAGIQLGIDGVDVYKSKWLENRNDTRARSVVVTVYALBELDPV 180
Db 107 LIRQOVTENTNTAIALLEGIGRGYRSTQQALETWLDNRNDARSIIILERVIALELDIT 166
Qy 181 AKIPSAISQGEVPLLSVYQAANLHLILLRDAISFGAEWGTPEGISTFYDRQVTRTAQ 240
Db 167 TAILPFIREEVPLLVYQAANLHLILLRDAISFGSEWGWASSDVNQYQEQRITPEE 226
Qy 241 YSDYCVKQWNTGLDKLKGNTAAASWLKTHQPREMTLLVLDLVALPNTDTRYPYETTAQ 300
Db 227 YSNHCVQWNTGLANLRCNTAESAELRYNQFRRDLTLGLVLDLVALPNTDTRYPYETTAQ 286
Qy 301 LTRRVYTDPIVFNRETSGPCRRWSLNSDISFSEVESAVIRSPHLFDLSLEFYTTTQAG 360
Db 287 LTRSIYTDPIGRTNAPSGFASTNWFNNAPSFAIEAAIPRPHLLDFPEQLTIYS--AS 344
Qy 361 LPLANNTEYLEYVGHGSIKYK-----NTNASSALERNYGTITSNKIKYDLANKDIFQVRS 416
Db 345 SRNSESQTQHMVYVGHRLNFRPIGGTLNTSTQGLTNTSINPVLQF---TSRDVYRTESN 401
Qy 417 GADLANYAQVGVYPYASFTLLDKNTGSGVGGTYSGEHTTMQVCTQNYNTIDRIPEE- 475
Db 402 AGTNILFTTPVNGVPMFARFNPQO-NIYERGAITYSQYOGVGI--QLFDSLELPPET 458
Qy 476 -NEPLSRGYSHRLSHITSYSPSKNASSPARVGN--LPVPANTHRSADVMTVYVSDKITQ 531
Db 459 TERPNESYSHRLSHIGLI-----IGNTLRAPVYNTWRSADRTWIGPNRITQ 507
Qy 532 IPVYKAITLVSGTIVIRKPGFTGGMILKRTSSGLPAYSVSVKSPLSQRYRARIYAST 591
Db 508 IPLVKALNLHSGVTVVGGPGTGGDILLRTWTGTGTGDIRLNLINVLSPSQRYRARIYAST 567
Qy 592 NLRPLVTSIGTRIYSINWNTKNGDDLTFTWPLATIGTAFTEPSNYSDSLTVGADSPAS 651
Db 568 DLQPFTRINGTIVNIGNFRSRTMNRGDNLEYSFRTAGFSTPFNFNAOSTFTLGHQSP-S 626
Qy 652 GGEVYVDKFEIIPVATFEAEEDLVAKKAVNGLFSTKDD-ALQTSVTDYQVNNQANLVE 710
Db 627 NQEVYIDRVFVPAEYFEAEYDLERAQKAVNALFTSTNPRLLKTDVTDYHLDQVSNWA 686
Qy 711 CLSDLEYPNKRLMDAVKAEKRLVQARNLQDTGKRLNG----- 751
Db 687 CLSDPECLDERELFEKVKYAKRLSDERNLQDPNFTFISGQLSPASIDGQSNPPSINEL 746

```


Db 621 GNGEYVVDKIELIPVATFAEYDLERAQAVNALFTNTPRRLKTDVTDYHIDQVSNLV 680
Qy 710 ECLSDLEYPNKRMMLDWAQKAKRLVQARNLLQDTPGNRIN----- 750
Db 681 ACLSDPECLDEKRELLEKVKYKRLSDERNLLQDPNFTSINKQDPFISTNQSNFTSIEH 740
Qy 751 -GNGWTGSGTGVAEQDVLKFDKSLRLTSAREIDTETPTLYQQIDESLLKYTRYKL 809
Db 741 QSEHGWSGSENIITQEGNDVFKENYVTLPGT---FNECYPLYQKIGESLKYTRYQL 797
Qy 810 KGFIGSSODLEIKLIRRANQIVKNVP--DNMLPDVL--PVNSCGGIDRCSEQQOYVDANL 865
Db 798 RGYLEDSDLEIYLRYNKAKETIDVQTESLWPSVESPIRCGEBNRC--PHFENP 855
Qy 866 ALENKENG-NMSSDSHAFSPHIDTGEIDLNENTGIVVFKIPTNGYATLGNLELVEBG 924
Db 856 DLDCSCRDGEKCAHSHHFLSDIDVGCTDLHENLGVVWPKIKTQEBGARLGNLEPIEEK 915
Qy 925 PLSETTERAQOQOQOQDQKARVKGASEKAYAAQAIIDLPAQYQDQKLSGVESDM 984
Db 916 PLLGEALSrvKRAEKWRDKEKLQLETKRKYVYTAKEAVDALFVDSQYDLQADNTIGMI 975
Qy 985 LAAQNLVQSPVYVNDALPEIPGNYTSFTLTNRLQAAWNLVLDLNAIPNGDFENGLSD 1044
Db 976 HAADKLVRIRRAYLSLPEIPGVNABIPBELEGHITATSLYDARVVRKGDFFNGLTC 1035
Qy 1045 WNATSDVNOQSDTSVLVIPNMNSQVSQFTVPQNYRVLRTARKEGVDGVYIIRDG 1104
Db 1036 WNVKGVHDVQSHRSDLVPEWEAEVSAVRVFCGCGYILRVYAYKEGVEGCVTHIEI 1095
Qy 1105 ANQETITLP----- 1113
Db 1096 ENNTDELKFRNREERYVPTDGTGNDYTAHQGTAGCADACNSRAGVEDAYEVDTASV 1155
Qy 1114 -----NICDDTAGVLSAD--QTSYITKTVEFTPEQWIDMSETEGV 1154
Db 1156 NYKTYEEYTYDRDNHCEYDRGYNPPVPAGYVTKLEYPEPTDTWIEIGETEGK 1215
Qy 1155 FNIESVELVLEER 1167
Db 1216 FIVDSVELLMEE 1228

RESULT 10

ID CIBC BACTM STANDARD; PRT; 1233 AA.
AC Q45774;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Pesticidal crystal protein cryIbC (insecticidal delta-endotoxin
DE CryIbC) (Crystalline entomocidal protoxin) (140 kDa crystal protein).
GN CryIbC OR CryIbC OR CryIbC.
OS Bacillus thuringiensis (subsp. morrisoni).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OK NCBI_taxid=1441;
RN [1]
RP SEQUENCE FROM N.A.
RA Bishop A.H., Bone E.J., Ellar D.J.;
RT "Cloning of novel Bacillus thuringiensis delta-endotoxin";
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF INSECTS.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; Z46442; CAAG6568.1; -
DR HSPF; P02965; IC1Y.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin C.
DR InterPro; IPR005639; endotoxin N.
DR InterPro; IPR008979; Gal bind like.
DR Pfam; PF00555; endotoxin_1.
DR Pfam; PF03944; endotoxin_C_1.
DR Pfam; PF03945; endotoxin_N_1.
KW Toxin; Sporulation.
SQ SEQUENCE 1233 AA; 140451 MW; 7318382413529F21 CRC64;

Query Match 41.2%; Score 2491; DB 1; Length 1233;

Best Local Similarity 42.1%; Pred. No. 3.4e-134; Indels 144; Gaps 20;
Matches 536; Conservative 208; Mismatches 384;

Qy 1 MSPNNQNEYILDASSSTVSNSRYRYPLANDQTTTLQNNMYXDYLMSGEENPELFCNP 60
Db 1 MTSNRKNEELINALSIPTVSNPSTQNLSPD-----ARIEDSLCVAR-----VNNI 47
Qy 61 ETPSSSTVQIGIGVQVGLGALGVPPAGQIATSTYFIVGQLMPSSTVSVMIMKQVED 120
Db 48 DPFSASTVQVGINAGRIILGVLPVPPAGQIATSTYFIVGQLMPSSTVSVMIMKQVED 106
Qy 121 LIDQKITDSVRKATAGLQGLDGLDVYQKSLKNLENRNDTRARSVVVTVIALLELDFV 180
Db 107 LIRQVTEHTNTAIALEGLEGVRSYQQALETWLDNENDARSIIILRYVALELDT 166
Qy 181 AKIPSPAISGOEVLVSYQAANLHLLLDASI FGAEWGFTTGEISTFTYDRQVTRAQ 240
Db 167 TAIPLFIRNEEVLVSYQAANLHLLLDASIFGSEMGMASSDVNVQYQEQIRYTEE 226
Qy 241 YSDYCVQVNTGLDKLGTNAASHLYHQPREMTLVLDLVALFPNYDVTYPIETTAQ 300
Db 227 YSHCVQVNTGLNLNLTGNABSLRYNQPRDLTGLVLDLVALFPNYDVTYPIETTAQ 286
Qy 301 LTRVYTDPIVFNRETSGGFCRRMSLNSDISFSVESAVIRSPHLFDLSIEFYTTAG 360
Db 287 LTRVYTDPIVFNRETSGGFCRRMSLNSDISFSVESAVIRSPHLFDLSIEFYTTAG 344
Qy 361 LPLANTEYLEYVWGHSGIKYK-----NTNASSALERNYGTITSNKIKYDLANKDFQVNSL 416
Db 345 SRWSSTQBMVYVWGHRLNFRPIGGLTNTSTQGLTNTSINPVTLOF---TSRDVYRTESN 401
Qy 417 GADLANYYAQVGVYPYASFTLLDKNTGSGVGGTYSKPHMTMQVCTQVNTIDEIPPE- 475
Db 402 AGTNILFTTPVGVVPFARFNPINQ-NIYERGATTSQYQGVGI--QLPDSLETLPET 458
Qy 476 -NEPLSRGYSHRLSHITSYSFSKNAASPARYGN---LPVFAWTHRSADVNTVYSDKITQ 531
Db 459 TERPNESYSHRLSHIGLI-----IGNTLRAPVYSWTHRSADRTWIGPNRITQ 507
Qy 532 IPVKAHTLVSGTTVIRKPGFTGNILKRTSSGFLATYSVSKGSLSORIARINASTT 591
Db 508 IPLVKALNLHSGVTWVGPGPGTGGDILRRTWGTFGDIRLNLNINVPLSQRYVRIRYASTT 567
Qy 592 NLRPLVITISGTRIVSINVNKTNKGGDILFTNTPLATIGTATFTSNYSDSLTVGADSPAS 651
Db 568 DLQPTFRINGTGTWVGNFSPRTMNRGDLNLEYSFRTAGSTFPNPLNAQSTFTLGAQSP-S 626
Qy 652 GGEVYVDFELIPVNAATPEAREEDLVAKVANGLFTSKD--ALQTSVTDYQVQNAANLVE 710
Db 627 NQEVYIIRVEVPVPAEVTFAEYDLERAQAVNALFTSTNPRRLKTDVTDYHIDQVSNVA 686
Qy 711 CLSDLYPNKRMMLDWAQKAKRLVQARNLLQDTPGNRIN----- 751
Db 687 CLSDPECLDEKRELLEKVKYKRLSDERNLLQDPNFTFISQLSPASIDGQSNFPIBEL 746

Search completed: June 21, 2004, 13:37:53
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:34:30 ; Search time 56 Seconds
(without alignments)
6575.174 Million cell updates/sec

Title: US-10-089-678-1
Perfect score: 6044
Sequence: 1 MSPNNQNEYILDASSSTSV.....MSTEGVFNISSVELVLEE 1167

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL_25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mbc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	5303.5	97.7	1144	2	Q8KZL7
2	2750.5	45.5	1144	2	Q45745
3	2545.5	42.1	1228	2	Q93T75
4	2539.5	42.0	1228	2	Q93NM5
5	2433	40.3	1231	2	Q8KNV2
6	2056.5	34.0	1155	2	Q9F296
7	2050	33.9	1180	2	Q9S5V8
8	2048	33.9	1176	2	Q7WZT9
9	2045	33.8	1176	2	Q45736
10	2034	33.7	1176	2	Q9RC30
11	2032.5	33.6	1155	2	Q93T21
12	2016.5	33.4	1171	2	Q06894
13	2003	33.1	1169	2	Q8GHE8
14	1997	33.0	1189	2	Q9L877
15	1989.5	32.9	1174	2	Q45749
16	1953.5	32.3	1160	2	Q93TF9

ALIGNMENTS

RESULT 1

Q8KZL7 Q8KZL7 PRELIMINARY; PRT; 1144 AA.

AC Q8KZL7; 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Cry8 protein.
GN Cry8.
OS Bacillus thuringiensis (subsp. galleriae).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=29338;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SDS-502;
RA Asano S., Yamamoto T.;
RT "a novel cry8 gene highly toxic to Anomala cuprea.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB089299; BAC07226.1; -
DR GO; GO:0015070; P:toxin activity; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005639; endotoxin C.
DR InterPro; IPR008979; Gal_bind_like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
SQ SEQUENCE 1144 AA; 128060 MW; 98F93070C49014AB CRC64;

Query March 97.7% Score 5903.5; DB 2; Length 1144;

Best Local Similarity 97.9%; Pred. No. 0;

Matches 1143; Conservative 0; Mismatches 1; Indels 23; Gaps 1;

Qy 1 MSPNNQNEYILDASSSTSVSDNSVRYPLANDQTTTLQNNYKYDYLRMSEGENPELFCNP 60

Db 1 MSPNNQNEYILDASSSTSVSDNSVRYPLANDQTTTLQNNYKYDYLRMSEGENPELFCNP 60

Qy 61 EFTFISSTVGTGIGVQVIGALGVFPFAGQIASPVFVQGLMPSSTSVSWEMIMKQVED 120

Qy		498	NASSPARYGNLPVANTHRSADVNTVYSKIIQIIPVKKAHTLVSGTIVIKGPGFTGGM	557
Db		480	NIYS---GNLPTVWTHRDVLTTITADRIITQLPLAVKSFEIPAGTVVVRGPFGTGDI	535
Qy		558	LKXTSSGPLATYSVSVKSPLSOBYRARIYASVTNLFLFVITISGRILIYSINVNKNANKD	617
Db		536	LRTGTGVGTOTIRVRTTAPITQBRYIRFRFPASTNLPFIGRVGDROVNPYDFGRTWNRGD	595
Qy		618	DLTFFNTFLATIGTAPTFSNYSDSLTVGADSPASGGEVYVDKFELIPVNAVTAPEABDLDV	677
Db		596	ELYESPAFREFTDFNFROPBELISVFANAFSAGQEVYDRLEIIPVPAREAKEDLEA	655
Qy		678	AKKANGLPFSKKDALQTSVTDYQVNOANLVECLSDLEYPNKRMIMDAVKEAKLVAQA	737
Db		656	AKKAVASLPTRTRDGLQVNVKDYOQVQAAANLVSCLSDBOQGYDKOMLLEAVRAAKELSRE	715
Qy		738	RNLLODTGFNRING--ENGWGTSGTIEVAGDVLFDORSRLITSAREIDTETPTLYXQQ	795
Db		716	RNLLODPDFNTINSTBENGWKASNGVVISREGPPFYKGREALQLASAR-----ENPTYTYQK	771
Qy		796	IDESLLKPTRYTKLGPIGSSODLEIKLIBRRANQIVKNVPDNLLDPLVPVNSCGGIDRC	855
Db		772	VDASELPKYTRYSDSGPVKSSQDLEILIHKKHVHLVKNVPDNLVSTPYDDSCSGINRC	831
Qy		856	SEQOYYVDANLALENWGE-NGNMSSDSHAFSHIDTGEIDLNENTGIWVVFKIPTTNGYAT	914
Db		832	OBOQWNAQLETHEHPMDCCBAAQTHERFSYIDTGDLSNSVDQGIWAIFKVRTTGYAT	891
Qy		915	LGNLELVBEGLPSGETLERAQOQEQQODKWARKRGASEKAYIAAKQAIIDLRFADYQDOQK	974
Db		892	LGNLELVVGFLPSGESLEREQDRONTKWSAEJGRKRAETDRVYQDAKQSINHLPDYQDOQK	951
Qy		975	INSGVEMSDMLAQNLIQSIPIYVYNBALPEIPGMNYTSPTELNRLQOANNLYDLRNAIP	1034
Db		952	LNPEIGHADIMAQNLVASISDVYSDAVLQIPGINYEIXYTELSNRLQOASLYLTSRNAVQ	1011
Qy		1035	NGDFRNGLSDWNATSDEVNVQOLSPTSVLVTPNNMSQVSQQOPTQPNRYRYLVRTARKEGV	1094
Db		1012	NGDFNGLSDSNATAGASVQDQGNTHFLVISHWDQVUSQQFVQPNCXYVLRTAEKVG	1071
Qy		1095	GDGYVIIRDGANOTETTLTFNICDDDTGVLSDQTSYITKTVEFTPSPTEQWIMDSFEGE	1154
Db		1072	GDGYVITRDGAHHETTTLTFNACDYDINGTVYVTONTYLTKEVIFYPSHTHEMMWEVNETEGA	1131
Qy		1155	FNTBSVLVLREE	1167
Db		1132	FHDISEFVETEK	1144

RESULT 3

```

Q93T75 PRELIMINARY;      PRT; 1228 AA.
AC Q93T75.
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Delta-endotoxin Cry1Ba2.
DE CRY1BA2.
OS Bacillus thuringiensis (subsp. entomocidus).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_taxID=1436;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HD-9;
RA Mat Isa M.N., Abdullah M.A.P., Mahadi N.M.;
RT "Characterization of cry1B gene and its flanking regions cloned from
RL Bacillus thuringiensis subsp. entomocidus HD-9";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF363025; AAK51084.1; -.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin C.
DR

```



```

532 IPVYKHTLVSGTIVKPGPTGGNHLKRTSSG-----PLAYTSVSVKSPLSOR 580
508 IPAVKGRFLFNG-SVIGSGPTGDDVRLNRNNGNIQNRGIEVPTQSTST-----R 560
581 YRARIYASTTNLRFLPTISGTRIYSINVNKTMMKGGDDLTPNTFDLATIGTAPTGSYSD 640
561 YRVRYASVTSIELNVNIGNSSIPTNTLPATAASLDNLQSGDPFGVLEINNAFTSATGN- 619
641 SLTVGADSFASGGEVYVDKFBELIPVNAITFEAEEDLDVAKCAVNGLFSTSKD-ALQTSVTD 699
620 --IVGARFSAEVIIDRFEPITATFEAYDLERAKAVNALFTSTNPRRLKTDVTD 677
700 YQVNOANLVECLSDLEYPNEKMLMDAVKAKRLVQARNLLQDTGPFNRING----- 751
678 YHIDQSNVAVACLSDDEFCLDEKRELFKVKYAKRLSDERNLLQDPNFTFISQQLSPASD 737
752 -----ENGWGTSGTGIKVAEGDVLKORSRLTSAREIDTETVPTVLYQOIDS 799
738 QQSNTFSINELSEHMGWSENVTIQEGNDVFKBNVTYLPGT---FNECYPNTLYOKIGES 794
800 LLKPTRYKLGKFGTSSQDLIELIRHRANQIVKNVP--DMLLPDVL--PVNSCGIDRC 855
795 ELKATRYQLAGYIEDSODLEIYLIRYNAKHETLDVPGTDSLWPLSVKSPICGCEPNRC 854
856 SEQYVDANLALENNGENG-NMSSDSHAFSPHIDTGEIDLNENTGIVVVPKIPPTNGYAT 914
855 A--PHFEWNPDLDCSCRDGERCAHSHHFTILDIDVGCTDLHENLGVWVVPKIKTOBGYAR 912
915 LGNLEVEBGLSGETLERAOQOQOQODKARKGSEKAYAAKQALDRLPADYQDOOK 974
913 LGNLEFIEBKPLIGEALSRVKRAKKMKHKKQLQETKRVVTEAKETVDALFVDSHYNR 972
975 LNSGVMSDMLAAQNLVQSIPVYVNDALPEIPGMNVTSTFELTNRLQOAWNLYLRLNAIP 1034
973 LQADTNIGMHADELVHRHIEAYLPPLPFGINAVIFEELENIHSTAFSLYDARNVIK 1032
1035 NGDFRNLGSDNNATSDVNVQQLSDTSUVI PNMNSQVSOQPTVQPNRYVLRVARTKGV 1094
1033 NGDFNGLSLCNVKGHVQVQSHRSRLVPEWEAEVSAQVAVRVCPCRGVILRVATYKGY 1092
1095 GDGVVIRDGANQTETLTENICDD-----DTGVLS-----A 1125
1093 GEGCVTHIEINNYDELKFKNCEBEVPTDTGTCNDTATAGTAAACNSRNAGYDAEYV 1152
1126 DQTS-----YITKVTFTSTEQVWIDM 1148
1153 DTTASVNVKPYEBETTYDVRDNHCEYDGVVNVPPVPAGVYTKLEVPETDTVWIEI 1212
1149 SETGVFNIESVELVLEEB 1167
1213 GETEGKFAVDSVELLMBE 1231

RESULT 6
Q9P296 ID Q9P296 PRELIMINARY; PRT; 1155 AA.
AC Q9P296;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Delta endotoxin.
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]
RP SEQUENCE FROM N.A.
RA Meza-Basso L.A., Theoduloz C.;
RT "Cloning and expression of a delta endotoxin gene from a Chilean
native Bacillus thuringiensis strain."
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U94191; AAG16877.1; -.
DR PIR; A29125; A29125.
DR HSSP; P02965; 1CIY.

```

DR GO:0015070; P:toxin activity; IEA.
 DR GO:0006952; P:defense response; IEA.
 DR InterPro; IPR001178; Endotoxin.
 DR InterPro; IPR005638; endotoxin C.
 DR InterPro; IPR005639; endotoxin N.
 DR InterPro; IPR008979; Gal_bind_like.
 DR Pfam; PF00555; endotoxin; 1.
 DR Pfam; PF03944; endotoxin C; 1.
 DR Pfam; PF03945; endotoxin N; 1.
 DR SQU SEQUENCE 1155 AA; 130557 MW; 5D69E3B2F527749D CRC64;

Query Match 34.0%; Score 2056.5; DB 2; Length 1155;
 Best Local Similarity 39.2%; Pred. No. 1.2e-108;
 Matches 477; Conservative 184; Mismatches 389; Indels 167; Gaps 29;

QY 53 NPGL-----FGNPET-FISSSTVQTG---IGI---VGQVLGALGVPPAGQIASFYSP 97
 DB 4 NPNINECPYCNLSNPEVEVLGGERIEFTGYPIDISLSLTQPLLSFVPGAGFVLGLVDI 63
 QY 98 IVGQLMPSTVSVMEMIMKQVEDLDKITSVRKTALAGLQGLDGLDVYQKSLNMLE 157
 DB 64 IWGIFGPSQ---WDAPLVQIEQLNQRIEFPARNQALSRLEGLSNLYQIYAESPREWEA 119
 QY 158 NRNDTRARSVVTVIALELDFVAKIPSPAGQEVPLLSVYQAANLHLLLDASIFG 217
 DB 120 DPTPALREERKIQNDMNSALTTPALPLFAVQNYQVPLLSVYQAANLHLSVLRDVSVPF 179
 QY 218 AEWGTPGEISTEFDYDQVTRTAQTSYDCKVNTGLDKLGTNAASLWKYHQFRREMTLL 277
 DB 180 QRWGFDAATINSRYNDLRLGNTYTHAVRWNTGLERVNGPDSRDHRYNQFRELTLT 239
 QY 278 VLDLVALPNYDTRTYPIETTAQLFREYTPPIVFNRETSGGCRWRSLNSDISFSRYES 337
 DB 240 VLDIVSLFPNYSRTYPIRTVSQLTREIYTPVLENFD--GSP--RGSAGQ-----IEG 289
 QY 338 AVIRSPHLPDILSEIEFTTACGLPLANTVLEYLVVGHSHK-----YKNTWASSALERNY 392
 DB 290 S-IRSPHMLDILNSITTYT-----DAHREGEY--YWSGHQIMASPVGFSGPBFTFPL---Y 338
 QY 393 GTITSNKIKYDLANKDIFQVRSICADLANVYAOVY--GVVPASFTLDDKNT---GSGS 446
 DB 339 GTMGNAAPQQRVIAQLGGVYRTLSSTL---YRFPNIGINNQQLSVLDTGEPAGTSSN 395
 QY 447 VGGFTYSKPTHTMQVCTQYNTIDIPENNE--PLSRGYSHRLSHITSY--SPSKNASSP 502
 DB 396 LPSAVYRKSQTV-----DSLDEIPPQNNVPPRQGFSHRLSHVSNFRSGFSNYSVSI 447
 QY 503 ARYGNLVPFANTHSADVTNTVYSDKITQIPVKAHTLVSGTWTVKPGFTGGMILKRTS 562
 DB 448 IR---APMPSMHRSAFPNNIIPSSQITQIPUTKSTNIGSGTSVVGKPGFTGGDLIRTS 504
 QY 563 SGPLAYTSSVSKSPLSQRYARIRYASTTNLRFLVTISGTRIYSINNVKTNMKGDLDLTFN 622
 DB 505 PQGISTLRVNTAPLSQRYRIRYASTTNLQFTSIDGRPINQGNFSATMSSGSLQSG 564
 QY 623 TFDLATICTAFTFNSYSLTVGADSPASGGEVYVDKFKELIPVNAATPABEDLDVAKQAV 682
 DB 565 SFRTVGFTTFPNSGSSVFTLSAHVFNSGNEVYIDRIEFVPAEVTFEARYDLERAQKAV 624
 QY 683 NGLFTSKD-ALQTSVTVQVNOANLVRCISDELYPNEKRMLDVAKKAEKLVQARNLL 741
 DB 625 NELFTSSNQIGLKTVDVTHIDQVSNLVECLSDPECLDEKSELSEKVKHAKRLSDERNLL 684
 QY 742 QDTGPNRLNG--ENGWGTSGTGEVAGDVLKDRSLRITSAREIDETPTLYLYQOIDES 799
 DB 685 QDPHLRGLNRQLDRGWRSTDIITQGGDDVFKENVYLLGTFD---ECYPTLYLQKIDES 741
 QY 800 LKPYTRYKLKGFSSQDLKILHRANQIVKQVNPDLPLVFNSSCGSIDRCSBQQ 859
 DB 742 KLKAYTRYQLRGYIEDSQDLLEYLIRYNAKHETVNVPGT--GSLWPLSAPSPIGKCAHH- 798
 QY 860 YVDANLALNNGENGNSSDSHAFSDHTDGTEDILNENTGIVWVPKIPITTYGATYLNLE 919

DB 799 -----SHFSLDIDVCGCTDLNSELGLVWVIFPKIKTQDGHARLGNLE 838
 QY 920 LVERGPLSGETLERAQOQOQODKMKRRGRGASEKAYTAAKQAIIDLRFADYQDKLNSGV 979
 DB 839 FLEEKPLVGEALARVKRAEKWRDKLEWETNIVYKEAKESVDALFINSQVDRLOADT 998
 QY 980 EMSDLAAQNLVQSIPIYVYNDALEIPICQNTYTSFTLNLRLQQAANLVDLRNAINPNDPR 1039
 DB 899 NIAMIHAADKRVHSIRREAYLPESLPIGVNAAIFEELEGRISTAFSLVDARNVIRKNGDFN 958
 QY 1040 NGLSDMNATSDVNV--QQLSDTSVLVIPNNNSQVSOQFTVQPNRYRVLRVTARKEGVGDGY 1098
 DB 959 NGLSCWNVKGVHDVBEQNHRSVLVPEWEAEVSEVVCBGRGYILAVTAYKEGVGEC 1018
 QY 1099 VIIRDGANQTELTAFN-----ICDDDTGV----- 1122
 DB 1019 VTIIHEIKNTDELAKFSNCVBEVFPNNTVTCNDYTATQBEYEGTYTSRRNRGYDGAYESNS 1078
 QY 1123 -LSADQTS-----YITKTVBFTSTEQVMDNSE 1150
 DB 1079 SVPADYASAYEKKAYTDGRRDNPCSNRGYGDYTPLPAGYVTKLELYFPETDKVWIRIGE 1138
 QY 1151 TEGVFNIESVELVLEEB 1167
 DB 1139 TECTFIVDSVLELLMEE 1155

RESULT 7
 Q9SSV8 PRELIMINARY; PRT; 1180 AA.
 AC Q9SSV8;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE BT84Al crystal protein (Crystal protein CryIA).
 GN BT84Al.
 OS Bacillus thuringiensis, and
 OS Bacillus thuringiensis (subsp. sotto).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1428, 29340;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=B.thuringiensis, STRAIN=T84Al;
 RA Nagamatsu Y.;
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=B.thuringiensis; STRAIN=T84Al;
 RA Ogo M., Yamada S., Kobayashi Y., Shibata J., Nagamatsu Y.;
 RT "Nucleotide sequence of the Lepidoptera-toxic Protein Gene of Bacillus
 thuringiensis subsp. dendrolimus T84Al.";
 RL J. Fac. Appl. Biol. Sci. Hiroshima Univ. 29:95-107(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=B.thuringiensis; STRAIN=T84Al;
 RA Nagamatsu Y., Imai Y., Hatanaka C., Funatsu G., Hayaishi K.;
 RT "A Toxic Fragment from the Entomocidal Crystal Protein of Bacillus
 thuringiensis.";
 RL Agric. Biol. Chem. 48:611-619(1984).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=B.thuringiensis (subsp. sotto);
 RA Zhong W.F., Cai P.Z., Yan W.Z., Zhang Z.X., Xiang Y.M.;
 RT "A cryIIA gene cloned from Bacillus thuringiensis serovar sotto.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB026261; BAA7213.1; --
 DR EMBL; AF510713; AAM44305.1; --
 DR PIR; A22798; A22798.
 DR HSSP; P02965; 1CIY.
 DR GO; GO:0015070; P:toxin activity; IEA.
 DR GO; GO:0006952; P:defense response; IEA.
 DR InterPro; IPR001178; Endotoxin.
 DR InterPro; IPR005638; endotoxin C.
 DR InterPro; IPR005639; endotoxin N.
 DR InterPro; IPR008979; Gal_bind_like.
 DR Pfam; PF00555; endotoxin; 1.
 DR Pfam; PF03944; endotoxin C; 1.
 DR Pfam; PF03945; endotoxin N; 1.
 DR SQU SEQUENCE 1155 AA; 130557 MW; 5D69E3B2F527749D CRC64;

InterPro: IPR005639; endotoxin N.
 DR InterPro: IPR008979; Gal bind like.
 DR Pfam: PF00555; endotoxin; 1.
 DR Pfam: PF03944; endotoxin C; 1.
 DR Pfam: PF03945; endotoxin N; 1.
 SQ SEQUENCE 1180 AA; 133489 MW; 1199BA6D1DCB62D CRC64;

Query Match 33.9%; Score 2050; DB 2; Length 1180;
 Best Local Similarity 39.3%; Pred. No. 3e-108;
 Matches 481; Conservative 174; Mismatches 414; Indels 156; Gaps 26;

QY 53 NPDL-----PCNPET-PISSSTVQTG---IGI---VCGVLTGALGVPPAGQIASFYSP 97
 DB 4 NPNINECIPYCNLSNPEVEVLGGERIETGTPIDISLSLTQFLSEFPVPGAGFVLGLVDI 63
 QY 98 IVGQLMPSSTVSVMEMIMKQVEDLIDQKITSVRKKTALAGLQGLDGLDVKYQSLKNWLE 157
 DB 64 IWGIFGPSQ---WDAPLVQIEQLINQRIEFPARNQAIISRLGSLNLYQIYASFRWEA 119
 QY 158 NRNDTRARSVVVVYQYIALHLDLFVAKIPSPAIISGQVEPLLSVYAQAAMHLHLLLRDASIFG 217
 DB 120 DPTNPALREBMRIOFNMDNSALTAIPLPAVQYQVPLLSVYVQAAMHLHSLVLRDVSFVG 179
 QY 218 AEWGFTPGHISTFYDQVTRTAQYSDYCVKWYNTGLDKLGTNAASWLKYHOFEREMTLL 277
 DB 180 QRWGFAAATINSRYNDLTRIGNYTDYAVRWYNTGLERVPDSDRWVRVYQFRELTLT 239
 QY 278 VLQDLVLPNYDTRTYPIETTAOLTRVYTDPIVFNRKTS-GGFCRWSLNSDISFSEVE 336
 DB 240 VLQDLVLPNYDTRTYPIETTAOLTRVYTDPIVFNRKTS-GGFCRWSLNSDISFSEVE 290
 QY 337 SAVIRSPHLPDILSIEFYFT-TRAGLPLANNTEYLEYVWGHSHK-----YKNTNASSALER 390
 DB 291 ---IROPHLMDILNRIITYDVRHG-----FNYSWGHIQITASVPVPGSPGFAPPLFG 339
 QY 391 NYGTTISNKIKYDLANKDIQVRS-----LGLDANIYAAQYGVVPYASFTLLDKNT 442
 DB 340 NAGNAAPVPL--VSLTGLGIFRTLSPLRYRRIILGSGPNQQLFVLDTGTFSPASLTNNL 397
 QY 443 GSGSVGGTGYSKHPTMOVCTQNTVIDETPEPNE--PLSRGYSRHLRSHITTSYSPSKNAS 500
 DB 398 PS-----TIYRQGTV-----DSLVDIIPQDQNSVPPRAGFSRHLRSHVITMLSQAGAV 444
 QY 501 SPARYGNLPVPAWTHRSADVTNTVYSDKITQIPVKAHTLVSGTTVIKPGFTGGMILKR 560
 DB 445 YTLR---APTFSWQRSAEENNIIPSSQITQIPLTKSTNLGSGTSVVKVGPGTGGDILR 501
 QY 561 TSSGFLAYTSVSKSPISQRYRARIYASITNLRLFTVTSIGTRIYSINVKNTWKNGDDLT 620
 DB 502 TSPQISTLRVNTAPLSQRYRARIYASITNLRLFTVTSIGTRIYSINVKNTWKNGDDLT 561
 QY 621 PNTFDLATICTAPTFSNYSDSLTVGADSPASGGEVYVDFKELIPVNAATFEABEDLOVAKK 680
 DB 562 SGSFRTVGTFTTTPNFNSGSSVFTLSAHVFNFSNGEVIYIDRIEFPVETVFEAYDLEPRAQK 621
 QY 681 AVNGLFTSKKD-ALQTSVTDYQVNAANLVECLSDLELTPNEKRLMDVAPKEAKELVQARN 739
 DB 622 AVNELFTSSNQIQLKTDVTDHYDQVSNLVECLSDLELTPNEKRLMDVAPKEAKELVQARN 681
 QY 740 LLQDTGPNRLNG--ENGWTSSTGTEIYVAGDVLFWDRSLRITSARBITDTEPTPYLVQOQID 797
 DB 682 LLQDPNFRGINRQLDRGWRGSDTIQGGDDVFKENYVTLIGTFD---ECYPTLVYQKID 738
 QY 798 ESLLKPTVKYKLGFGSSODLETKLIRHARQIVKXVPDN---LLPDLVLPVNSCGGID 853
 DB 739 ESLLKATRYQLRGYEDSDLETKLIRHARQIVKXVPDN---LLPDLVLPVNSCGGID 798
 QY 854 RCSQQVVDANLALENNGENG-NMSSDSHAFPHIDTGEIDLNENTGIWVVFKIPTNGY 912
 DB 799 RCA--PHLEWNPDLDCRGEKCAHSHFSLDIDVGCTDMLNEDLGVWIFPKIKTQDGH 856
 QY 913 ATTLGNLLEVBEGPLSGTELRACQEQOQWDRMKRKGSAKAYYAAKQAIIDLFLADYQD 972

RESULT 8
 Q7WZT9
 ID Q7WZT9 PRELIMINARY; PRT; 1176 AA.
 AC Q7WZT9;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Delta-endotoxin.
 GN CRYLAA.
 OS *Bacillus thuringiensis*.
 OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; *Bacillus*.
 ON NCBI_TaxID=1428;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ly30;
 RA Yao J., Zhang J., Song F., Chen Z., Li C., Huang D.;
 RT "A novel cryIIA gene from Bt strain ly30."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP384211; AAP80146.1;
 SQ SEQUENCE 1176 AA; 133169 MW; 725C14EA940B869B CRC64;

Query Match 33.9%; Score 2048; DB 2; Length 1176;
 Best Local Similarity 39.3%; Pred. No. 3.8e-108;
 Matches 480; Conservative 174; Mismatches 414; Indels 152; Gaps 26;

QY 53 NPDL-----PCNPET-PISSSTVQTG---IGI---VCGVLTGALGVPPAGQIASFYSP 97
 DB 4 NPNINECIPYCNLSNPEVEVLGGERIETGTPIDISLSLTQFLSEFPVPGAGFVLGLVDI 63
 QY 98 IVGQLMPSSTVSVMEMIMKQVEDLIDQKITSVRKKTALAGLQGLDGLDVKYQSLKNWLE 157
 DB 64 IWGIFGPSQ---WDAPLVQIEQLINQRIEFPARNQAIISRLGSLNLYQIYASFRWEA 119
 QY 158 NRNDTRARSVVVVYQYIALHLDLFVAKIPSPAIISGQVEPLLSVYAQAAMHLHLLLRDASIFG 217
 DB 120 DPTNPALREBMRIOFNMDNSALTAIPLPAVQYQVPLLSVYVQAAMHLHSLVLRDVSFVG 179
 QY 218 AEWGFTPGHISTFYDQVTRTAQYSDYCVKWYNTGLDKLGTNAASWLKYHOFEREMTLL 277
 DB 180 QRWGFAAATINSRYNDLTRIGNYTDYAVRWYNTGLERVPDSDRWVRVYQFRELTLT 239
 QY 278 VLQDLVLPNYDTRTYPIETTAOLTRVYTDPIVFNRKTS-GGFCRWSLNSDISFSEVE 336
 DB 240 VLQDLVLPNYDTRTYPIETTAOLTRVYTDPIVFNRKTS-GGFCRWSLNSDISFSEVE 290
 QY 337 SAVIRSPHLPDILSIEFYFT-TRAGLPLANNTEYLEYVWGHSHK-----YKNTNASSALER 390
 DB 291 ---IROPHLMDILNRIITYDVRHG-----FNYSWGHIQITASVPVPGSPGFAPPLFG 339
 QY 391 NYGTTISNKIKYDLANKDIQVRS-----LGLDANIYAAQYGVVPYASFTLLDKNT 442

```
Db 340 NAGNAAPPVL--VSLTGLGIFRTSSPLRYRIILGSGPNNQELFVLDGTBFSFASLTNNL 397
Qy 443 GSGSVGGPTYSKPHTTMQVCTQNTYIDIPPE--PLSRGYSHRLSHITSYSPSKNAS 500
Db 398 PS-----TIYRQRGT-----DSLVDVFPQDMSVPPRAGPSHRLSHVMTLSQAAGAV 444
Qy 501 SPARYGNLPVPAWTHRSADVTNTVYSKDTQIPVVKAAHTLVSGTIVVKGPGTGGNILKR 560
Db 445 YTLR---APTFSQWHSRAEFNIIIPSSQITQIPLTSTNLAGSGTSVVGKPGPTGCDILRR 501
Qy 561 TSSGLATYSVSKPSLQSRVRYASTTNLRFPVITSGTRISYNNVTKNGDDLT 620
Db 502 TSPGOISTLRVNTAPLSQRVRYASTTNLRFPVITSGTRISYNNVTKNGDDLT 561
Qy 621 FNTDLATIGTAFTSNYSDSLTVGADSFASGGEVYVDKFLIPVNAFPAEEDLDVAKK 680
Db 562 SGRFRTVGTPTFPNFSNGSVPTLSAHVPSNGEVYIDRFPVPAETPEAYDLERAKQ 621
Qy 681 AVNGLFTSKKD-ALQTSVTDVQVNAANLVECLSDLYPNRKMMDAVKRAKRLVQARN 739
Db 622 AVNELFTSSNITGLKTDVTDYHIDQVSNLVECLSDLYPNRKMMDAVKRAKRLVQARN 681
Qy 740 LLQDTGENRING--RNGWTSGTIEVARGDVLFKDRLSLRLTSAREIDTETPTLYQQID 797
Db 682 LLQDPNFRGNRQLDGRWGSDDITIQGDDVFKENYVTLGTFD---ECPYTYLYQKID 738
Qy 798 ESKLKYRYKLGPIGSGQDLLEIKLIRHRANQIVKNVDPN----LLPDVLPVNSCGGID 853
Db 739 ESKLKYRYKLGPIGSGQDLLEIKLIRHRANQIVKNVDPN----LLPDVLPVNSCGGID 798
Qy 854 RCSBOQVYDANLALNNGENG-NMSSDSHAFSHIDTGEIDLNENGTWVFKIPTTNGY 912
Db 799 RCA--PHLEWNPDLDCSCRDGKCAHSHHSFLDIDVGCTDLNEDLGVVWVIFKIKTDQGH 856
Qy 913 ATGLNLELVESGTPSLGETLERAQOQOOWOMARKERKASAKYAAQADRLPADYOD 972
Db 857 ARGLNLELVESGTPSLGETLERAQOQOOWOMARKERKASAKYAAQADRLPADYOD 916
Qy 973 OKLNGSVMSDKLAQNLVQSIPIPVYNDALPEIPGMNTYTSFELTNRLQQAQWNLVLENA 1032
Db 917 DQLOADTWIAMIHAADKRVHSIREAYLPVLSVIFGVNAAIPEELGRIFTAFLYDAENV 976
Qy 1033 IPNGDFRNLSDNATSDVNV--QQLSDTSVLVIPNNNSQVQOFTVQPNRYVLRVATRK 1091
Db 977 IKNGDFRNLSDNATSDVNV--QQLSDTSVLVIPNNNSQVQOFTVQPNRYVLRVATRK 1036
Qy 1092 EGVGDGYVIIRDGANQETITFN-----ICDD-----1118
Db 1037 EGVGDGYVIIRDGANQETITFN-----ICDD-----1118
Qy 1119 DTGVLSDQTS-----YITKTVEFTPSTQWID 1147
Db 1097 EAPSPADYASVYEKSYDGERNPFCHNRGYDYTPLPVGYVTKLEYFPETDKWIB 1156
Qy 1148 MSETGVNIESVELVLEE 1167
Db 1157 ICETEGTPIVDSVELLAMEE 1176
```

RESULT 9

```
Q45736
ID Q45736 PRELIMINARY; PRT; 1176 AA.
AC Q45736;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Insecticidal crystal protein.
GN CRYIA(A).
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=FU-2-7;
RA Kondo S., Tamura N., Kunitate A., Hattori M., Akashi A., Ohmori I.;
RT "Cloning and nucleotide sequencing of two insecticidal delta-endotoxin
RL genes from Bacillus thuringiensis var. kurstaki HD-1 DNA.";
RW Agric. Biol. Chem. 51:455-463(1987).
RP SEQUENCE FROM N.A.
RC STRAIN=FU-2-7;
RA Udayasuriyan V., Nakamura A., Mori H., Masaki H., Uozumi T.;
RT "Cloning of a new cryIA(a) gene from Bacillus thuringiensis strain
RL FU-2-7 and analysis of chimaeric CryIA(a) proteins for toxicity.";
RW Biosci. Biotechnol. Biochem. 58:830-835(1994).
DR EMBL; D17518; BAA04468.1; -.
DR PIR; JC2219; JC2219.
DR HS9P; P02965; 1CIV.
DR GO; GO:0015070; P:toxin activity; IEA.
DR GO; GO:0006952; P:defence response; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin C.
DR InterPro; IPR005639; endotoxin N.
DR InterPro; IPR008979; Gal bind_like.
DR Pfam; PF00555; endotoxin_1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
DR CONFLICT 77 77 L -> P (IN REF. 1).
DR CONFLICT 965 965 S -> P (IN REF. 1).
DR PT CONFLICT 965 965 S -> P (IN REF. 1).
SQ SEQUENCE 1176 AA; 133075 MW; AA113594A289P85 CRC64;

Query Match 33.8%; Score 2045; DB 2; Length 1176;
Best Local Similarity 39.3%; Pred. No. 5.7e-108;
Matches 479; Conservative 174; Mismatches 415; Indels 152; Gaps 26;

Qy 53 NPGL-----FGNPET-FISSSTVQGT---IGI---VGQVLGALGVFPAGQIASFYSP 97
Db 4 NPNNECIPYCNLSNPEVEVLGGERIETGYTPIDISLSTQTLSEFPVFGAGFVLGLVDI 63
Qy 98 IVGQLWPSSTVSVMIMKQVEDLIDQKITSVKRTALAGLQGLGDLGVOKSLRWLE 157
Db 64 IWGIPGPSQ---MDAFLVQIEQLINRIEFPARNOAISRLGLSNLYQIYAESFWEA 119
Qy 158 NRNDTRARSVVVQYIALEDFVAKIPSPALSGQEVPLSVYQAANLHLALLLRASIFG 217
Db 120 DETNPALREERIOPDNMSALTTAIPLLAVQNYQVPLSVYQAANLHLVLRDVSFVG 179
Qy 218 AEWGTPGEISTFYDRQVTRTAQYSDYCVKWTGCLDKLGTNAASWLYKHQFRREMTLL 277
Db 180 QRMGDAATINSRYNDLRLIGNTYDYAVRWNTSLERWGPDSRDVRYNQFRELTLT 239
Qy 278 VLDLVALFPNYDTRTYPIETTAQLTRVYTDPIVFNRETS--GGFCRRWSLNSDISFSRVE 336
Db 240 VLDLVALFPNSVDSRRYPIRTVSQLTRVYTNPLENFDGSPRGMARIEQN-----290
Qy 337 SAVTESPHLPIILSEIEFT--TRAGLPLNNTEYLEVWVCHSIK-----YKNTNASSALER 390
Db 291 ---IRPHLMDILNSITITVDVHRG-----FNTWSGHQITASPVGSGGEFAPPLPG 339
Qy 391 NYGTTISNKIKYDILANKDIFQVRS-----LGADLANYAYQVYGVYASFTLLDKQNT 442
Db 340 NAGNAAPPVL--VSLTGLGIFRTSSPLRYRIILGSGPNNQELFVLDGTBFSFASLTNNL 397
Qy 443 GSGSVGGPTYSKPHTTMQVCTQNTYIDIPPE--PLSRGYSHRLSHITSYSPSKNAS 500
Db 398 PS-----TIYRQRGT-----DSLVDVFPQDMSVPPRAGPSHRLSHVMTLSQAAGAV 444
Qy 501 SPARYGNLPVPAWTHRSADVTNTVYSKDTQIPVVKAAHTLVSGTIVVKGPGTGGNILKR 560
Db 445 YTLR---APTFSQWHSRAEFNIIIPSSQITQIPLTSTNLAGSGTSVVGKPGPTGCDILRR 501
Qy 561 TSSGLATYSVSKPSLQSRVRYASTTNLRFPVITSGTRISYNNVTKNGDDLT 620
Db 502 TSPGOISTLRVNTAPLSQRVRYASTTNLRFPVITSGTRISYNNVTKNGDDLT 561
```

```
Qy 621 ENTFLATICTAFTFSNYSDSLTVGADSPASGGVYVDKPELIPVNATFEABEDLDVAKK 680
Db 562 SGSPRTVGTTPPFPNSGSSVFTLSAHVFNSEVNIIDRIEFVPAEVTFBAEYDLERAQK 621
Qy 681 AVNGLFTSKKD-ALQTSVTDYQVNOAANLVECLSDLEYLPNEKRLMDAVKAKELVQARN 739
Db 622 AVNELFTSSNQIGLKTVDVTHIDQVSNLVECLSDLEKQSEKVKHAKLSDBRN 681
Qy 740 LLODTGPNRING--ENGWTSSTGIEVAGDVLFDKRSRLTSAREIDTETPTLYQKID 797
Db 682 LLODPNFRGINRQLDRGWSGTDITIOGGDDVFKNVYVTLTGTFD---ECYPTLYQKID 738
Qy 798 ESLLKPYTRYKLGFGISSQDLLEIKLIRHRANOIVKQVDPN-----LLPDVLVFNSSCGGID 853
Db 739 ESLLKPYTRYKLGFGISSQDLLEIKLIRHRANOIVKQVDPN-----LLPDVLVFNSSCGGID 798
Qy 854 RCSEQQVVDANLALENNENG-NNSSDSHAPSFIHDTGCEIDLNTENTGIWVFKIPTNGY 912
Db 799 RCA--PHLEWNPDLDCSCRDGEKCAHSHHPSLDIDVGCTDLNHDLGWVVIKIKTQDGH 856
Qy 913 ATIGNLELVEEGLSGETLERAOQOQWODKMARKGASKEYAAKQADRLFPADYQD 972
Db 857 ARGNLEFLBEKPLVGEALARKVREKREKLEWETNIVYKEAKESVDALFVNSQY 916
Qy 973 QKLSGVEMSDMLAAQNLVOSIPYVNDALPEIFGNYTSTFELTNLQOAWNLIDLNA 1032
Db 917 DQLOQDTNIAHIAHADKRVHSIREAYLPESLVIKQVNAIPEELBEGRAISTAFSLYDARNV 976
Qy 1033 IPNGDFRGLSDMNTSDVNV-QQLSDTVLVIWNWSQVSOQFTVPQNTRYLVRTARK 1091
Db 977 IKNGDFRGLSCWNVKGVHDVEEQNQRSVLVWPEAEVSEVQVRVCPGRTYILRVAYK 1036
Qy 1092 EGVGDGVVIRDXGANQETETFN-----ICDD-----1118
Db 1037 EGYEGCVTTHIEIENNTDELKFSVCVEEBIYPPNVTVCNDTTVNQEYGGAYTSRNGYN 1096
Qy 1119 DTGVLSDOTS-----YITKTVEFTPTSEQWID 1147
Db 1097 EAPSVADYASVYEKSYTDGRRENPCENFRNGRYDTPPLPGVYVTKLEYFPETDKWIE 1156
Qy 1148 MSETGVPNTIESVELVEE 1167
Db 1157 IGETEGTFIVDSVELLME 1176

RESULT 10
Q9RC30 PRELIMINARY; PRT; 1176 AA.
AC Q9RC30;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 135 kDa insecticidal protein.
OS Bacillus thuringiensis (subsp. kurstaki).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_Taxid=29339;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HD-1-02;
RA Hou B.K., Chen Z.H.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF154676; AAD55382.1; -.
DR HSP; P02965; ICIY.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin C.
DR InterPro; IPR005639; endotoxin N.
DR InterPro; IPR008979; Gal_bind_like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin C; 1.
DR Pfam; PF03945; endotoxin N; 1.
```

```
SQ SEQUENCE 1176 AA; 133010 MW; FCBEO69D0B081D8C4 CRC64;
Query Match 33.7%; Score 2034; DB 2; Length 1176;
Best Local Similarity 39.2%; Pred. No. 2.4e-107;
Matches 478; Conservative 173; Mismatches 417; Indels 152; Gaps 26;

Qy 53 NPFL-----FCNPET-FISSSTVQVG---IGI---VGQVLGALGVPVPAQGIASFYSF 97
Db 4 NPINECIPYNCLSNPEVEVLGGRIETGYTPIDISLSTQFLSEFVGAGFVLGLVDI 63
Qy 98 IVGQLWPSSTVSVWEMTMKQVEDLIDOKITDSVRKTAGLQGLGDGLDVYOKSLKNWLE 157
Db 64 IWGIFGFSQ-----WDTFLVQIEQLINRIEFPARNQALSRLEGSLNLIQIYABSFREWA 119
Qy 158 NRNDTRARSVVVTVYIALELDFVAKIPSAISFALSGOEVPLLSVYQAANLHLLLRDASIFG 217
Db 120 DPTNPALREEMRIQFNDMNSALTTAIPLLAVQNYQVPLLSVYQAANLHLSVLVDVSFVG 179
Qy 218 AEWGFTPEEISTFVDQRVTRTAQYSDYCVKWTNTGLDKLKTNAASWLKYHQFREMTLL 277
Db 180 QRMGFDAAATINSRYNDLTRIGNTYDVAVRWYNTGLERVMGPDSDRDVRYNQFRELTUT 239
Qy 278 VLDLVALFPNVDTYTPLETTAQLTRVYTDPIVFNRETS--GGPCRRWSLNSDISFSFVE 336
Db 240 VLDLVALFNSYDSRRYPIRTVSQLTREYTNVPLENFDGSPFGMAQRIEQN-----290
Qy 337 SAVIRSPLHFDLSIEISFTT--TRAGLPLNNTLEYLWVGHSHIK-----YKNTNASSALER 390
Db 291 ---IRQPHLMDLINSIITYDVHSG-----FNWMSGHQUITASPVGSPGPEAPPLFG 339
Qy 391 NYGFTSNKI KYDLANKDIPQVRS-----LGADLANYYAQVGVYPYASFLLDKRT 442
Db 340 MAGNAAPVVL--VSLTGLGIFRTLSSPLRYRRIILGSGPNQBELFVLDGTFEFSALTTNL 397
Qy 443 GSGSVGGFTYKPHWTWQVCTQNTYNTIDEIPENE--PLSRGYSHLSHITSYSFKNAS 500
Db 398 PS-----TIYRQGTV-----DSLDPVTPQONSVPFRAGFSHRLGHVTHLSQAGAV 444
Qy 501 SPARYGNLPFAWTHRSADVNTVYSDKITQIPVVKAAHTLVSGTTVIKGPFTGGNILKR 560
Db 445 YTLR--APTFSWQHRSAEFNNIIPSSQITQIPLTKSTNLGSGTSVVKGPGTGGDILRR 501
Qy 561 TSSGPLAYTVSVKSPLSQRYRARIYASTNRLFLVTISGTRIYSINVKNTMKGDDLT 620
Db 502 TSPQIISTLRVNIAPLSQRYRARIYASTNRLFLVTISGTRIYSINVKNTMKGDDLT 561
Qy 621 ENTFLATICTAFTFSNYSDSLTVGADSPASGGVYVDKPELIPVNATFEABEDLDVAKK 680
Db 562 SGSPRTVGTTPPFPNSGSSVFTLSAHVFNSEVNIIDRIEFVPAEVTFBAEYDLERAQK 621
Qy 681 AVNGLFTSKKD-ALQTSVTDYQVNOAANLVECLSDLEYLPNEKRLMDAVKAKELVQARN 739
Db 622 AVNELFTSSNQIGLKTVDVTHIDQVSNLVECLSDLEKQSEKVKHAKLSDBRN 681
Qy 740 LLODTGPNRING--ENGWTSSTGIEVAGDVLFDKRSRLTSAREIDTETPTLYQKID 797
Db 682 LLODPNFRGINRQLDRGWSGTDITIOGGDDVFKNVYVTLTGTFD---ECYPTLYQKID 738
Qy 798 ESLLKPYTRYKLGFGISSQDLLEIKLIRHRANOIVKQVDPN-----LLPDVLVFNSSCGGID 853
Db 739 ESLLKPYTRYKLGFGISSQDLLEIKLIRHRANOIVKQVDPN-----LLPDVLVFNSSCGGID 798
Qy 854 RCSEQQVVDANLALENNENG-NNSSDSHAPSFIHDTGCEIDLNTENTGIWVFKIPTNGY 912
Db 799 RCA--PHLEWNPDLDCSCRDGEKCAHSHHPSLDIDVGCTDLNHDLGWVVIKIKTQDGH 856
Qy 913 ATIGNLELVEEGLSGETLERAOQOQWODKMARKGASKEYAAKQADRLFPADYQD 972
Db 857 ARGNLEFLBEKPLVGEALARKVREKREKLEWETNIVYKEAKESVDALFVNSQY 916
Qy 973 QKLSGVEMSDMLAAQNLVOSIPYVNDALPEIFGNYTSTFELTNLQOAWNLIDLNA 1032
Db 917 DQLOQDTNIAHIAHADKRVHSIREAYLPESLVIKQVNAIPEELBEGRAISTAFSLYDARNV 976
```



```
QY 1033 IPNGDFRNLGSDWATSDVNV-QQLSDTSVLVLPNNNSOVSOQFTVQPNRYVLRVARTAK 1091
DB 1034 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 977 IKNGDFNGLSCWVKGHVDVEQNNQRSLVVPWEAEVSEVVRVCPGEGVILRVATYK 1036
QY 1092 EGVGQGVYIIRGANOTETLTN-----ICDD-----1118
DB 1037 EGTGGCGVTHIEHNTDLSKPSNCVEEIIYNNVTTCNDYTNQBEYGVATSNRGN 1096
QY 1119 DTGVLSDQTS-----YITKVFETPSTEQVWD 1147
DB 1097 EAPSVADYASYVEEKSYTDGRRENPCFNRGVRDYTPLPVGVYKLELYPETDKVWIE 1156
QY 1148 MSTREGVNIESVELVLEE 1167
DB 1157 IGETEGTFTVDSVELLME 1176

RESULT 11
Q93721
AC Q93721; PRELIMINARY; PRT; 1155 AA.
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Crystal protein CryIAb16.
GN CryIAB16.
OS Bacillus thuringiensis (subsp. israelensis).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1430;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=407;
RC Yu J., Tan L., Wu D., Pang Y.;
RT "Molecular characterization of a silent gene encoding a 130-kilodalton
RT crystal protein from Bacillus thuringiensis subsp. israelensis.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF375608; AAK5546.1; -.
DR FIR; A29125; A29125.
DR GO; GO:0015070; P:toxin activity; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal_bind_like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
SQ SEQUENCE 1155 AA; 130747 MW; 7F0C98E0100C7698 CRC64;

Query Match 33.6%; Score 2032.5; DB 2; Length 1155;
Best Local Similarity 38.9%; Pred. No. 2.9e-107;
Matches 473; Conservative 185; Mismatches 392; Indels 167; Gaps 29;

QY 53 NPEL-----FGNPET-FISSVTQGT---IGI---VGQVLGALGVFPAGQIASFYSF 97
DB 4 NPNINECIPYNCLSNEVEVLGGERIETGYTPIIDISLSTQFLSEFVFGAGFVLGVDI 63
QY 98 IVGQLWPSVTSVWEMIMKQVRLDIDOKLTDVSRKTALAGLOGLDVYOKSLKNWLE 157
DB 64 IWGIFGPSQ-----WDAFLVQIQLINQRIEFPARNQAIISRLGSLNLYQIYAESFWEA 119
QY 158 NENDTRARSVVVTVQIALBELDFVAKIPSPAISQEVPLLSVYQAANLHLLLRASIFG 217
DB 120 DPTNPALREEMRIQFNWNSALTATPLFAVQNYRVPLLSVYVQAVNLHLSVLRLVLF 179
QY 218 ASWGTFPGISITFTRQVTRTAQSYCVKWNNTGLDKLGTNAASLWKHGFRRMTLL 277
DB 180 QRWGFDAATINSRYNDLTRIGNYTDHVRWYNTGLERVWGPDSRDWIRYQFRRLTUT 239
QY 278 VLDIALPNNYDTRTPYIETTAQLTEVYTDPIVENFTSGGFCRWLSLNSDISFEVES 337
DB 240 VLDIVSLFPNYSRTYPIETVTSQLTREIYTNPVLENFD--GSF--RGSAGQ-----IRG 289
```

```
QY 338 AVIRSPHLEPILSEIEFTTTRAGLPLANNTEYLEYVWVCHSIK-----YKNTNASSALERNY 392
DB 290 S-IRSPHLEPILSEIEFTTTRAGLPLANNTEYLEYVWVCHSIK-----YKNTNASSALERNY 398
QY 393 GTITSNKIKYDILANKDIFQVRSIGADLANYYAQVY--GVFYASFLLDKNT---GSGS 446
DB 339 GTWGNAAPOORIVAQLGQGVYRLSTL---YRPPNIGINNQQSLVDLDTFAYGTSSN 395
QY 447 VGGFTYSKPHHTMQVCTQNYNTIDEIPPEB--PLSRGYSHRLSHITSY--SPSKNASSP 502
DB 396 LPSAVTRKSGTV-----DSLDRIPPQNNVPRQGFSHRLSHVSMPSGFSNSVSI 447
QY 503 ARYGNLPVPAWTHRSADVNTVYSKIQIPIVKAHTLVSGTIVVIGPGPTGNGILKRTS 562
DB 448 IR---APMFSWIHRSABFNIIIPSSQITQIPLTKSTNLGSGTSVWVGPGTGGDILRRTS 504
QY 563 SGPLAVTSVSKSPLSQRYRARIYASTNLRFLVTISGTRILYSINVKNTMKGDDLTN 622
DB 505 PQOISTLRVNTIAPLSQRYRARIYASTNLRFLVTISGTRILYSINVKNTMKGDDLTN 564
QY 623 TFDLATIGTAFPSYSDSLTVGADSFASGGEVYVDKFLIPVNTATFEAEEDLDVAKAV 682
DB 565 SRTVGFTTTPFNSGSSVFTLSAHVFNSENVYIDRIEFPVPAEVTFEAEYDLERAQKAV 624
QY 683 NGLFTSKD-ALQTSVTDYQVNOAAMLVESLSDLYPNKRMWDVAKRAKRLVQARNLL 741
DB 625 NELFTSSNQIGLKTVDVTHIDQVSNLVECLSDFCLEKELSKVYKRAKRLSDERNLL 684
QY 742 QDTGNRING--ENGWGTSGIEVARGDVLFKDRSLRITSAREIDTETPTLYVQQLDES 799
DB 685 QDPNFRGINRQDRGWRGSTDITIQGGDDVFKENYVTLTGTFD---ECYPTLYQKIDES 741
QY 800 LKPYTRYLKGFGIGSSQDLKILKIRHRANQIVKXNVDPNLLPDLVLPVNSCGGIDRCSEBQ 859
DB 742 KLAYTRYQLRGYIEDSDLEIYLIRYNKAKHETVNVPGT--GSLWPLSAPSPIGKCAHH- 798
QY 860 YVDANLALNNGENGNMSSDSHAPSFIHDTGEIDLNENTGIWVVKIPITNGYATLGNLE 919
DB 799 -----SHHFLSDIDVGCTDLNEDLGVWVIFIKITQDCHARGLNLE 838
QY 920 LYEBGSLGETLERAQOQBOQODKWKARKGSEKAYAAKOAIIDLFLADYQDQKLSGV 979
DB 839 FLEEKPLVGEALARKVRAEKWRDKRKLWETNIVYKAKBSVDVAVFVNSQDRLQADT 898
QY 980 EKMSDLAAQMLVQSIPYVYNDALPRIPGNNYTSFTSLTNLQQAANLYDLRNAIPNGDPR 1039
DB 899 NIAMIHADKRVHSIREAYLPESLVTPGVNAALFELEGRIPTAFSLYDARNVKNKGDPN 958
QY 1040 NGLSDNATSDVNV-QQLSDTSVLVLPNNNSOVSOQFTVQPNRYVLRVARTKEGVGDGY 1098
DB 959 NGLSCWNVKGVHDVEEQNNHRSVLVVPWEAEVSEVVRVCPGEGVILRVATYKBEYGBGC 1018
QY 1099 VLIIRGANOTETLTN-----ICDDTGV-----1122
DB 1019 VTIIHEIENNTABLKPSNCVEEIIYNNVTTCNDYTNQBEYGVATSNRGN 1078
QY 1123 -LSADQTS-----YITKVFETPSTEQVWD 1150
DB 1079 SVPADYASAYEKAYTDGRDNPCESNRGVGYDTPLPAGVYVTRRELEVPPETDKVWIEIG 1138
QY 1151 TEGVNIESVELVLEE 1167
DB 1139 TEGTFTVDSVELLME 1155

RESULT 12
O06894
AC O06894; PRELIMINARY; PRT; 1171 AA.
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```


Db 305 NNLTIFDWFSGVRN-----FYWGCHRVISLIGGHNITSPYGRBANQBPSPFTNGP 359
Qy 410 IFQVRSAGADLANYYAQVGYVPYASFTLLDKNTGSGSGVGGFTYKSPHTMQVCTQNTYNTI 469
Db 360 VP--RTLSPNTRLRLQPPWAPPAPNLGRVGVGFSTPTNSFTYGRGTG-----DSL 409
Qy 470 DEIPPENE--PLSRGYSHRLSHITSYSPKSNASSPARYGNLPVPAWTHRSADVNTVYSYD 527
Db 410 TELPPEONSPPREGYSHRLCHAT---FVORSGETPFLTTGV-VFSWTHRSATLNTIDPE 465
Qy 528 KITQIPVVKAHVLTGTVTIGPGFTGKNTLSSGFLAVTYSVKSPLSQRVRRARY 587
Db 466 RINQIPLVKGRVWGTSVITGPGFTGDLIRNTFGDFVSLQVINSPIQTRRLRPY 525
Qy 588 ASTTNRLRFVT-----ISGTRIYSINVNTKMGDDLTNTFDLATIGTAFTFSYSD 640
Db 526 ASSRDARVIVLTGAAGTGVGVGVSVNMPLOKMTBEIGENLTSTRTFYTDPSNPPSFRANPD 585
Qy 641 SLTV-----GADSFASGGEVYVDPKPHLIPVATFEABEDLDVAKKAVNGLFTSKD-AL 693
Db 586 IIGISERPLFGAGISS--GELYIDKIHILADATFEABESDLERAQKAVNALFTSSNQIGL 644
Qy 694 QTSVTDYOVNQANLVCLSDLYPNKSKMLWDVAKKRLVQARNLLQDTGENRIN--G 751
Db 645 KTDVTDYHIDQVSNLVCLSDLYPNKSKMLWDVAKKRLVQARNLLQDTGENRIN--G 751
Qy 752 ENGWTGSGTIEVARGDVLFDKRSRLTSARIDTETPTLYQOIDSLLKPYRYKLKG 811
Db 705 DRGWRGSTDITQGGDDVFKENVYTLPGT--VD-ECYPTLYQKIDSKLKAYTRYBLRG 761
Qy 812 FIGSQOLEIKLIRHRANQIVNVPDN-----LLPDVLPVNSCGGIDRCSEQQYVDANLAL 867
Db 762 YIEDSQLEIYLIRYNAKHEIVNVEGTGSLWPLSAQSPIGRCGBPNRCA--PHLEWPNFL 819
Qy 868 ENNGENG--NMSDGHAFSPHIDTGEIDNENTGIVVVKIPTTNGYATLGNLELVESGPL 926
Db 820 DCSRDGEKKAHSHHFTLDIDVGCTDLNEDLGWVWVIFKIKTQGHAKLGNLEFLEKPL 879
Qy 927 SGETLEAAQQOQOQWQDMARKGRGASKAYYAAQAIIDRLPADYQOQKINSGVMSDLA 986
Db 880 LGEALAVKRAEKWRDREKQLETNIVYKEAKESVDALPVNSQYDRLQVDTMIAMHA 939
Qy 987 AQNLVQSIPIVYNDALEIPGQNTSFTLNTLRQQANLVLDLRLNAPNGDFRGLGSDWN 1046
Db 940 ADKRVHRIEAYLPESLVPICVNAAIPEELBGRIFTAYSLYDARNVIRKNGDFNGLLCWN 999
Qy 1047 ATSDVNV--QQLSDTSVLVIPNNNSQVQFTVQPNRYVLTAVTARKGVGDGVYVIRDGA 1105
Db 1000 VKGHVDVEQNNHRSVLVIPPEAEVSOEVRVCPGRGYILRVYATYKBYGEGCVTIHIE 1059
Qy 1106 NOTELTFN-----ICDDDTGV-----LSADQT 1128
Db 1060 DNTDELAFSCVBEVYPNNTVTCNNYTGTOBEYEGTYSSENQGYDRAFGNPNVPADYA 1119
Qy 1129 S-----YITKTVETPPTSPTEQWIDMSFEGVPI 1157
Db 1120 SVYEKSQYDGRNCPESNRGYDTPPLPAGYVTKOLEYFPETDKVWIEIGETGTFIV 1179
Qy 1158 ESVELVLEEE 1167
Db 1180 DSVELVLEEE 1189

RESULT 15
Q45749
ID Q45749 PRELIMINARY; PRT; 1174 AA.
AC Q45749;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Crystal protein (Crystal delta-endotoxin).
GN CRY1PB.
OS *Bacillus thuringiensis*.

Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
NCBI_taxID=1428;

OC NCBI_taxID=1428;
RN SEQUENCE FROM N.A.
RP STRAIN=BTS00349A;
RA Lambert B.;
RT "NO INFORMATION";
RL Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.
RN SEQUENCE FROM N.A.

RP STRAIN=B-Pr-88;
RA Li C., Zhang J., Huang D., Li G.;
RT "A crystal endotoxin from *Bacillus thuringiensis* strain B-Pr-88";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z22512; CAAB0235.1; -;
DR EMBL; AF336114; AA013295.1; -;
DR PIR; S32649; S32649.
DR HSSP; P02965; ICIV.

DR GO; GO:0015070; P:toxin activity; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin C.
DR InterPro; IPR005639; endotoxin N.
DR InterPro; IPR008979; Gal_bind_Like.

DR Pfam; PF00555; endotoxin_1.
DR Pfam; PF03944; endotoxin_C_1.
DR Pfam; PF03945; endotoxin_N_1.
SQ SEQUENCE 1174 AA; 133350 MM; 8C7F123F9446P15C CRC64;

Query Match 32.9%; Score 1989.5; DB 2; Length 1174;
Best Local Similarity 38.0%; Pred. No. 8.4e-105; Gaps 25;
Matches 458; Conservative 191; Mismatches 413; Indels 143;

Qy 59 NPETFISSSTVQTG-----IGIVGVGLGALGVPPAGQIASFYSPFIVGQLWPSSTVSVW 111
Db 17 NPEVILLSEESTGRLPLDISLSLTRLFLSFP-VPGVGVAFLGLDINGFITPSE---W 71
Qy 112 EMIMKQVEDLIDQKITDSVRKTALAGLQGLDLYQKSLKNWLENRNDTRARSVVVWQ 171
Db 72 SLFLQIQEQLIEQRIETLERNAITTLGLADSVEYLEARLBEENPNNAQLREBVRIR 131
Qy 172 YIALELDPVAKIPSPFASISGQEVPLLSVYAAANLHLLLRDASIPGAEWGFTPGHISTFY 231
Db 132 FANTDDALITAINNLTIFEFIPLLSVTVQAANLHLSLRDAVSFGQWGLDIATVNNHY 191
Qy 232 DRQVTRTAQYSDYCVKWTNTGLDKLKGNTAASWLKYHQFREMTLLLDLVALFPNYDTR 291
Db 192 NLLNLHRYTEHCLDTYNOGLENLRGNTTQWRSFNQFRRELTLTLDIVALPENYDAR 251
Qy 292 TYPIETTAQLPRVYTPDIVFNRETSGGFCRRMSLNSDISFSEVESAVIRSHPLFDLSE 351
Db 252 AYPQTSSQLTRFIYTSVIEDSPVSA-----NIPNGFNRAEFGV-RPPLHADPMNS 302
Qy 352 IEFYTRAGLPLANNTEYLEVWGHISIKYKNTNASSALERNYGTIT-----SNKIKY 402
Db 303 L-FVTA-----FTVRSQTWGHGLVSRNTAGNPINPPIYGFNPGGAIWATDSDPRPF 355
Qy 403 YDLANKDIFQVRSIGADLANYYAQVGYVPYASFTLLDKNTGSGSGVGGFTYKSPHTMQVC 462
Db 356 YRTLSDPVP-VRG-GFGNPHVVLGLGVAF-----QQTGTN-----HTRTPN 396
Qy 463 TQNTYNTIDEIPPENEPLS--RGYSHRLSHITSYSPKSNASSPARYGNLPVPAWTHRSADV 520
Db 397 SGTIDSLDEIPPDQNSGAPMWDYSHVLNHNVTFRWPKGEIAGSDSW-RAPMFSWTHRSADR 455
Qy 521 TNTVYSDKIQTOIPVVKAHVLTGTVTIGPGFTGKNTLSSGFLAVTYSVKSPLSQR 580
Db 456 TINIINPNIITQIPAVKKAHLHSGSTVVRGPGFTGDLRLRNTGTGADIRVNIITPFSOR 515
Qy 581 YRARIYASTTNRLRFVTISGTRIYSINVNTKMGDDLTNTFDLATIGTAFTFSYSD 640
Db 516 YRVRIYASTTDLQFFTRINGTSVNOGNFQRTMNRGNLSEGNPRTAGTSPFSPSNAQS 575

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:30:40 ; Search time 65 Seconds
(without alignments)
5072.814 Million cell updates/sec

Title: US-10-089-678-1
Perfect score: 6044
Sequence: 1 MSPNNQNEYILDASSTSV.....MSSTGVFNIESVHLEEE 1167

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: Geneseqp19808:*

2: Geneseqp19908:*

3: Geneseqp20008:*

4: Geneseqp20018:*

5: Geneseqp20028:*

6: Geneseqp20038:*

7: Geneseqp20038:*

8: Geneseqp20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6044	100.0	1167	5	Aau80281 Bacillus
2	4008	66.3	1157	2	Aar28900 Toxin 50C
3	4007	66.3	1157	2	Aar25997 Delta-end
4	4007	66.3	1157	2	Aar33768 Bt isolat
5	4007	66.3	1157	2	Aaw06418 Autiscara
6	3995	66.1	1157	2	Aar27343 B.thuring
7	3994	66.1	1157	2	Aar44208 Bacillus
8	3662.5	60.6	1149	2	Aar32354 Coleopter
9	3662.5	60.6	1149	2	Aar31692 B.thuring
10	3263.5	54.0	1210	5	Aau99256 Bacillus
11	3246.5	53.7	1206	5	Aau99255 Bacillus
12	3152	52.2	1169	2	Aaw06417 Autiscara
13	3128	51.8	1157	2	Aaw84581 Amino aci
14	3128	51.8	1157	2	Aaw84584 Amino aci
15	3126	51.7	1157	2	Aaw84571 Amino aci
16	3125	51.7	1157	2	Aaw84591 Amino aci
17	3125	51.7	1157	2	Aaw84575 Amino aci
18	3124	51.7	1157	2	Aaw84586 Amino aci
19	3124	51.7	1157	2	Aaw84582 Amino aci
20	3124	51.7	1157	2	Aaw84593 Amino aci
21	3124	51.7	1157	2	Aaw84588 Amino aci
22	3124	51.7	1157	2	Aaw84578 Amino aci
23	3124	51.7	1157	2	Aaw84576 Amino aci
24	3123	51.7	1157	2	Aar48678 Insectici
25	3123	51.7	1157	2	Aaw84570 Amino aci

26	3123	51.7	1157	2	Aaw84566 Amino aci
27	3123	51.7	1157	2	Aaw84569 Amino aci
28	3123	51.7	1157	2	Aaw84568 Amino aci
29	3123	51.7	1157	4	Aab82178 Cry9C #1.
30	3122	51.7	1157	2	Aaw84583 Amino aci
31	3122	51.7	1157	2	Aaw84592 Amino aci
32	3121	51.6	1157	2	Aaw84572 Amino aci
33	3121	51.6	1157	2	Aaw84577 Amino aci
34	3120	51.6	1157	2	Aaw84590 Amino aci
35	3120	51.6	1157	2	Aaw84579 Amino aci
36	3119	51.6	1157	2	Aaw84589 Amino aci
37	3118	51.6	1157	2	Aaw84587 Amino aci
38	3117	51.6	1157	2	Aaw84573 Amino aci
39	3115	51.5	1157	2	Aaw84567 Amino aci
40	3115	51.5	1157	2	Aaw84585 Amino aci
41	3114	51.5	1157	2	Aaw84580 Amino aci
42	3109	51.4	1157	2	Aaw84574 Amino aci
43	3100.5	51.3	1156	2	Aaw46857 Bacillus
44	3100.5	51.3	1156	2	Aay24960 Bacillus
45	3100.5	51.3	1156	4	Aau02034 B. thurin

ALIGNMENTS

RESULT 1
AAU80281
ID AAU80281 standard; protein; 1167 AA.
XX AAU80281;
XX
XX 30-JUL-2002 (first entry)
DT
XX
XX Bacillus thuringiensis insecticidal protein.
DE
XX
XX Insecticide; transgenic; Coleoptera larvae.
XX
XX Bacillus thuringiensis.
OS
XX
XX JP2002045186-A.
PN
XX
XX 12-FEB-2002.
PD
XX
XX 03-AUG-2000; 2000JP-00236140.
PF
XX
XX 03-AUG-2000; 2000JP-00236140.
PR
XX
XX (SDSB-) SDS BIOTECH CORP.
PA
XX
XX WPI; 2002-356468/39.
DR
XX
XX A protein having insecticidal activity, a DNA encoding said protein, and an agent and a method for preventing harmful organisms.
PT
XX
XX Claim 1; Page 9-12; 19pp; Japanese.
PS
XX
XX This invention relates to a crystalline protein comprising a fully defined sequence and the nucleotide sequence encoding this protein. The protein of the invention is an agent for preventing harmful organisms comprising Bacillus thuringiensis serovar galleriae SDS502, its mutant or a microbe transformed by a DNA encoding the protein. This microbe can be used to produce a protein containing the protein, or containing a protein having insecticidal activity produced by the SDS502, its mutant or a transformed microbe, a microbe which is transformed by using the above DNA and produces the above protein having insecticidal activity, a plant or a seed transformed by using the above DNA, and Bacillus thuringiensis serovar galleriae SDS502 producing a protein comprising and producing a protein showing insecticidal activity. The protein of the invention may have insecticidal activity. The agent is used for preventing Coleoptera larvae. This sequence represents the Bacillus thuringiensis insecticide protein of the invention
XX Sequence 1167 AA;
SQ


```
Query Match      100.0%; Score 6044; DB 5; Length 1167;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPNNQNEVEIILDASSSTSVSDNSVRYPLANDQTTLLQNNMYKDYLRMSGEGNPELFGNP 60
DB 1 MSPNNQNEVEIILDASSSTSVSDNSVRYPLANDQTTLLQNNMYKDYLRMSGEGNPELFGNP 60

QY 61 ETPFISSTVQTGIGIVGVLGALGVPFAGQIASFYSTFVGQLWPSSTVSVWEMIMKQVED 120
DB 61 ETPFISSTVQTGIGIVGVLGALGVPFAGQIASFYSTFVGQLWPSSTVSVWEMIMKQVED 120

QY 121 LIDQKITDSVRKTAAGLGLDGLDVYOKSLKNWLENRNDTRARSVVVTVYIALELDFV 180
DB 121 LIDQKITDSVRKTAAGLGLDGLDVYOKSLKNWLENRNDTRARSVVVTVYIALELDFV 180

QY 181 AKIPSPALSGQVPLLSVYAAQANLHLILRRDASIPGAEWGFTGCEISTFYDRQVTRTAQ 240
DB 181 AKIPSPALSGQVPLLSVYAAQANLHLILRRDASIPGAEWGFTGCEISTFYDRQVTRTAQ 240

QY 241 YSDYCVKMYNTGLDKLKGNAASWMLKYHOFREMTLLVLDLVALPNDYDTRTYPIETTAQ 300
DB 241 YSDYCVKMYNTGLDKLKGNAASWMLKYHOFREMTLLVLDLVALPNDYDTRTYPIETTAQ 300

QY 301 LTRVYVTDPIVNRRTSGGFCRRWSLNSDISPSEVESAVIRSPHLFDILSEIEFTYTRAG 360
DB 301 LTRVYVTDPIVNRRTSGGFCRRWSLNSDISPSEVESAVIRSPHLFDILSEIEFTYTRAG 360

QY 361 LPLANTEYLEYVWGHSHIKYKNTNASSALERNYGTITSNKIKYDYLANKDIPOVRSIGADL 420
DB 361 LPLANTEYLEYVWGHSHIKYKNTNASSALERNYGTITSNKIKYDYLANKDIPOVRSIGADL 420

QY 421 ANYAQVGVFPVASTLLDKNTGSGSVGGFTYSKPHTTMQVCTQNYNTIDEIPPENEPLS 480
DB 421 ANYAQVGVFPVASTLLDKNTGSGSVGGFTYSKPHTTMQVCTQNYNTIDEIPPENEPLS 480

QY 481 RGYSHRLSHITSYSPSKNASSPARVGNLPVFAWTHRSADVNTVYSDKIQIPVYKAHTL 540
DB 481 RGYSHRLSHITSYSPSKNASSPARVGNLPVFAWTHRSADVNTVYSDKIQIPVYKAHTL 540

QY 541 VSGTIVIKGPGTGNILKRTSSGPLAYTSVSVKPLSORVBARIRYASTTNLRLPVITS 600
DB 541 VSGTIVIKGPGTGNILKRTSSGPLAYTSVSVKPLSORVBARIRYASTTNLRLPVITS 600

QY 601 GTRIYSINVKNTMKGDDLTFTNTFLATIGTAFPTSNYSDSLTVGADSPASGGEVYVDKF 660
DB 601 GTRIYSINVKNTMKGDDLTFTNTFLATIGTAFPTSNYSDSLTVGADSPASGGEVYVDKF 660

QY 661 ELIPVNATPEAEEDLDVAKKANVGLFTSKKDALQTSVTDYQVQANLVCECLSDLYPNE 720
DB 661 ELIPVNATPEAEEDLDVAKKANVGLFTSKKDALQTSVTDYQVQANLVCECLSDLYPNE 720

QY 721 KEMLDWAKEARLVQARNLLODTPGNRNGENGWGTSGTIEVAGDVLFPKDRSLRTSA 780
DB 721 KEMLDWAKEARLVQARNLLODTPGNRNGENGWGTSGTIEVAGDVLFPKDRSLRTSA 780

QY 781 REIDTETPTLYLQOIDEILLKPYTKLKGFIGSSQDLIEIKLIHRANQIVKNVFDNLL 840
DB 781 REIDTETPTLYLQOIDEILLKPYTKLKGFIGSSQDLIEIKLIHRANQIVKNVFDNLL 840

QY 841 PDVLVNSCGGIDRCSEQQYVDANTALENNGENGMSSDSHAFSPHIDTGEIDLNENTGI 900
DB 841 PDVLVNSCGGIDRCSEQQYVDANTALENNGENGMSSDSHAFSPHIDTGEIDLNENTGI 900

QY 901 WVVKIPITNGYATLGNELVEEGLPSGETLBRACQOQOOWDWARKEGASEKAYIAAK 960
DB 901 WVVKIPITNGYATLGNELVEEGLPSGETLBRACQOQOOWDWARKEGASEKAYIAAK 960

QY 961 QAIDRLFADYQOQKLSNGVEMSDMLAAQNLVOSIPVYNDALPEIPGMNYTSTFELTNRL 1020
DB 961 QAIDRLFADYQOQKLSNGVEMSDMLAAQNLVOSIPVYNDALPEIPGMNYTSTFELTNRL 1020

QY 1021 QQAMNLYDLERNAI PNGDFRNLGLSDNNATS DVNQQLSDTSLVLPVHNSQVSOQFTVQPN 1080
DB 1021 QQAMNLYDLERNAI PNGDFRNLGLSDNNATS DVNQQLSDTSLVLPVHNSQVSOQFTVQPN 1080

QY 1081 VRYLVRVTARKEGVDGVYIIRDGANQTEITLTFNICDDDTGVLSDADQTSYITKTVFTPS 1140
DB 1081 VRYLVRVTARKEGVDGVYIIRDGANQTEITLTFNICDDDTGVLSDADQTSYITKTVFTPS 1140

QY 1141 TEQWIDMSSETGVFNIESVELVLEE 1167
DB 1141 TEQWIDMSSETGVFNIESVELVLEE 1167

RESULT 2
AAR28900
ID AAR28900 standard; protein; 1157 AA.
XX
AC AAR28900;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 01-APR-1993 (first entry)
XX
DE Toxin 50C.
XX
KW Endotoxin; acarides; pest; Two Spotted Spider; mite; phytophagus.
XX
OS Bacillus thuringiensis; kumamotoensis PS50C.
XX
PN W09219106-A1.
XX
PD 12-NOV-1992.
XX
PF 30-APR-1992; 92WO-US003546.
XX
PR 30-APR-1991; 91US-00693210.
PR 13-SEP-1991; 91US-00759248.
PR 30-SEP-1991; 91US-00768141.
XX
PA (MYCO ) MYCOGEN CORP.
XX
PI Payne JM, Cannon RJC, Bagley AL;
XX
WPI; 1992-398411/48.
DR N-PSDB; AAQ30821.
XX
PT New Bacillus thuringiensis isolates and toxins - used for controlling acarid pests of livestock, fowl, stored prods. and plants.
XX
PS Claim 16; Page 53 + 49-51; 62pp; English.
XX
CC Gene sequences encoding a toxin which is active against acarides and is obtainable from B. thuringiensis isolates PS17a, PS17b, 33f2, PS2A1, PS6SD1, PS6A1 and PS50C are given in AAQ30803-07 and AAQ30820-21 respectively. The toxin is a delta-endotoxin active against acarid pests, including the Two Spotted Spider mite. The isolates can be used against non-phytophagus mites such as acarid pests of livestock, fowl and stored prods. The genes can be cloned and used to transform other hosts, which can be used to control mites, or in the case of transgenic plants, be CC resistant to mites. (Updated on 25-MAR-2003 to correct PN field.) CC (Updated on 24-OCT-2003 to standardise OS field)
SQ Sequence 1157 AA;
Query Match 66.3%; Score 4008; DB 2; Length 1157;
Best Local Similarity 66.1%; Pred. No. 1.7e-255;
Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;
QY 1 MSPNNQNEVEIILDASSSTSVSDNSVRYPLANDQTTLLQNNMYKDYLRMSGEGNPELFGNP 60
DB 1 MSPNNQNEVEIILDASSSTSVSDNSVRYPLANDQTTLLQNNMYKDYLRMSGEGNPELFGNP 60
QY 61 ETPFISSTVQTGIGIVGVLGALGVPFAGQIASFYSTFVGQLWPSSTVSVWEMIMKQVED 120
```

Db 61 ETFFISSSTIQTGIGIVGRILGALGVFASQIASFYFVIGQLWPSKSDVINGEIMERVE 120
Qy 121 LIDQKITDSVRKLTALAGIGLGVQYKSLKNWLENNDTRARVVVVVTOYIALDLDFV 180
Db 121 LVQDKIEKYVKOKALAEGLGNALGVYQOSLEDWLENNDARTSRVSNQFALDLNFV 180
Qy 181 AKIPSPAISGQSVPLLSVVAQAANLHLLLRDASIFGAEWGFTPGGIBSTFYDQVTRTAQ 240
Db 181 SSIPSPAVSGHEVLLAVVAQAANLHLLLRDASIFGEWGFPTPGIBSRFYNRQVQLTAB 240
Qy 241 YSDYCVKNTGLDKGNALASWLYKHQFRREMTLLVLDLVALFNYDTRTFPIETTAQ 300
Db 241 YSDYCVKNTGLDKGNALASWLYKHQFRREMTLLVLDLVALFNYDTRTFPIETTAQ 300
Qy 301 LTRVETDPIVFNRETSGFCRWSLNSDISPSEVESAVIRSPHLPDILSEIRFYTRAG 360
Db 301 LTRVETDPIVFNRETSGFCRWSLNSDISPSEVESAVIRSPHLPDILSEIRFYTRAG 360
Qy 361 LPLNTEYLEYVWGHISIKYKNT-NASSALERNYGTITSNKIKYDILANKDIPOVRSGLAD 419
Db 361 ITLANDAYINYSGHTLYKRRRTADSVTVYANYGRITSEK-NSPALERDRIPEINSTVAN 419
Qy 420 LANYTAQVGVYASFTLIDKNTGSGVGGFTYSKPHHTMQVCTQNYNTIDEIPPEYE-P 478
Db 420 LANYTQKAYGVPGSWFMYKR--GTSSTTAYLYSKTHLTALQCTQVYESSDEIPLDRTVP 477
Qy 479 LSRGVSHRLSHITSYSPKASSPARYGNLPVPAWTHRSADVNTNTVYSDKIQIPVVKAH 538
Db 478 VASYSHRLSHITSYSPKASSPARYGNLPVPAWTHRSADVNTNTVYSDKIQIPVVKAH 538
Qy 539 TLVSGTIVIKGPGFTGNILKRTSSGFLAYTSVSKSPISQRYRARIYASTNTLALFTY 598
Db 536 MLYLGGSVVQPGFTGDIILKRTNPSILCTFAVTVNGSLSQRYRARIYASTNDEP-FTL 594
Qy 599 ISGTRIYSINVKNTMKGDLFTNTPDLATIGTAPFSPYNSLSLTVGADSPASGGEVYVD 658
Db 595 YLGDITIEKRNPKNDNGASLTAYETPKFASFTDFQFRETQDKILLSMGDFSSGQEVYD 654
Qy 659 KPELIPVNAATPEARDLDVAKAVNGLFTSKDALQTSVTDYOVNQAANLVECLSDLYP 718
Db 655 RLEFIPVDITYAEQDLAEAKAVNALFTNTKDLRPGVTDYEVNQAANLVECLSDLYP 714
Qy 719 NEKRLMDAVKEAKRLVQARNLLQDTGFNRINGENGTGSTEGLVAGDVLFKDRSLRIT 778
Db 715 NEKRLMDAVKEAKRLVQARNLLQDTGFNRINGENGTGSTEGLVAGDVLFKDRSLRIT 774
Qy 779 SARBIDTETPTLYQIDESLLKPYTRYKLGPIGSSQDLLEIKLIRHANQIVKXNPON 838
Db 775 GAREIDTETPTLYQIDESLLKPYTRYKLGPIGSSQDLLEIKLIRHANQIVKXNPON 834
Qy 839 LLPDVLVNSCGGIDRCSSQVVDANLALNNGENGNMSSDASHPSFHDITGEIDLNENT 898
Db 835 LLPDVLVNSCGGIDRCSSQVVDANLALNNGENGNMSSDASHPSFHDITGEIDLNENT 898
Qy 899 GIWVVFKEIPTNGYATLGNLELVEBGLSGETLERAQOQEQWQDKMARKRASKAYYA 958
Db 889 GIWVGFKITDPGSIATLGNLELVEBGLSGDALERLQREEQWQDKMARKRASKAYYA 948
Qy 959 AKQALDRLPADYQDQKLNKGVEMSLAAQNLVQSIPIPVYNDALPEIPGANTYSPTELTN 1018
Db 949 SKQAVDRLYADYQDQKLNKGVEMSLAAQNLVQSIPIPVYNDALPEIPGANTYSPTELTN 1008
Qy 1019 RLOQAWNLVDLNAIPNGDFRNLGNLSDNATSDVNVQOLSDTSVLVTPNWSQVSOQFTVO 1078
Db 1009 RLOQAWNLVDLNAIPNGDFRNLGNLSDNATSDVNVQOLSDTSVLVTPNWSQVSOQFTVO 1068
Qy 1079 PNRYVLRVTAKEGVGQGVYLRDGNANOTETLTTFMCDTGVLSADQTSYITKTVET 1138
Db 1069 PNRYVLRVTAKEGVGQGVYLRDGNANOTETLTTFMCDTGVLSADQTSYITKTVET 1128
Qy 1139 PSTEQWIDMSTEGVFNTIESVELVLEE 1167

Db 1129 PYTDOMWIEISETEGTFTYIESVELIVDVE 1157
RESULT 3
AAR25997
ID AAR25997 standard; protein; 1157 AA.
XX AAR25997;
XX AC AC
XX 25-MAR-2003 (revised)
XX 27-JAN-1993 (first entry)
XX Delta-endotoxin.
XX B.t.PS50C; colorado potato beetle; Leptinotarea decemlineata;
XX Pseudomonas; expression vector.
XX Bacillus thuringiensis.
XX EP498537-A2.
XX PN 12-AUG-1992.
XX PD 16-JAN-1992; 92EP-00300366.
XX PF 16-JAN-1991; 91US-00642112.
XX PR (MYCO) MYCOGEN CORP.
XX PA Poncerrada L, Sick AJ, Payne JM;
XX PI WPI; 1992-270497/33.
XX DR P-PSDB; AAR25997.
XX DR New insecticidal Bacillus thuringiensis strain NRRL B-18746 - and toxin,
XX PT Coleoptera.
XX PT encoding it, vectors and transformed cells, effective against
XX Claim 6; Page 11-14; 16pp; English.
XX The sequence given is an approx. 130 kD protien known as Bacillus
XX thuringiensis PS50C (B.t.PS50C). This novel delta-endotoxin has been
XX shown to be active against the Colorado potato beetle (Leptinotarea
XX decemlineata). The gene encoding this toxin can be transformed into
XX suitable hosts such as Pseudomonas, which can then be applied to the
XX environment of coleopteran insects where they will proliferate and be
XX ingested by the insects. This DNA sequence can be inserted into an
XX expression vector which contains a promoter/operator region, a ribosome
XX binding site, polyadenylation signals, etc. This will allow transcription
XX and translation of the cells in the appropriate host. (Updated on 25-MAR-
XX 2003 to correct PN field.)
XX SQ Sequence 1157 AA;
Query Match 66.3%; Score 4007; DB 2; Length 1157;
Best Local Similarity 66.1%; Pred. No. 1.9e-255;
Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;
Qy 1 MSPNNQNEVEILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSGRENPELPGNP 60
Db 1 MSPNNQNEVEILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSGRENPELPGNP 60
Qy 61 ETPFISSTVQGTGIGIVGVGLGVFPAGQIASFYFVIGQLWPSSTVSVWEMIMKQVED 120
Db 61 ETPFISSTVQGTGIGIVGVGLGVFPAGQIASFYFVIGQLWPSSTVSVWEMIMKQVED 120
Qy 121 LIDQKITDSVRKLTALAGIGLGVQYKSLKNWLENNDTRARVVVVVTOYIALDLDFV 180
Db 121 LVQDKIEKYVKOKALAEGLGNALGVYQOSLEDWLENNDARTSRVSNQFALDLNFV 180
Qy 181 AKIPSPAISGQSVPLLSVVAQAANLHLLLRDASIFGAEWGFTPGGIBSTFYDQVTRTAQ 240
Db 181 SSIPSPAVSGHEVLLAVVAQAANLHLLLRDASIFGEWGFPTPGIBSRFYNRQVQLTAB 240

420	LANYQAKYGVPGSWFPMVKR--GTSSTTHAYLSKXHTIALQGGCTQVYESSDBELPDRTPV	477
479	LSRGYSHRLSHITSYSESKNASSPARYGNLFPFAWTHRSADVNTVYSDKITQIPVVKAH	538
478	VABSYSHRLSHITSHSPSKNGS--AYYGSFPVFWHTTHSADLNTIYSDKITQIPAVKGD	535
539	TLVSGTTVIYKGPQFTGGNIIKRTSSGPLATYSVSKPSLQRTARIRIYASTNNLRILFVT	598
536	MLYLGSSVWGQPGFTGCDILKRTNPSILGTFAVTVNGSLSQRYRVRIRIYASTTDFB--FTL	594
599	ISGTRIYSINVNKTMKGGDLTNTPLDLAGTIGATPESNYSDSLTCGACDSFASGGEVVVD	658
595	YLGDTIEKRNFKMTMDNGASLTYYETFKPASFITDQFPRQTKILLSSMGDFSSQGEVYID	654
659	KPELIPVNATFEAEEDLDVAKKAVNGLFTSKDQALQTSVDYQVNOQANLVECLSDBLYP	718
655	RIFEPVDEIYEAEODLEAKKAVNALFTYTKGLRPGVTDYEVNOQANLVECLSDDLYP	714
719	NEKRLMDVAKEAKRLVOARNLLQDTGFNRINGENGWGTSGTIEVABEGDVLFKDRSLRLT	778
715	NEKRLLPDAVREAKRLSGARNLLQDPQFQINGENGWAASTGIEIVEGDVAPFKGYRLRLP	774
779	SAREIDTETPTVLYQOIDEISLAKPYTRYKLGKFIGSSODLEIKLIRHRANOIVKVPDN	838
775	GAREIDTETPTVLYQKVEEGVLKPYTRYILRGVFGSSQGLEIYTIIRHNTNRIVKVPDD	834
839	LLPDVLVNSCGGIDRCSEQYVDANALENNNGNGNMSDSHAFSPHIDTGHIDLMENT	898
835	LLPDVSPVNSDGSINRCSBKQYVNSRL-----EGENRSGDAHEFSLPIDIGELDYNEA	898
899	GIWVVFKIPITTYNGATLGNLELVEEGPLSGHETTLRAQQEQQWODKMARKEGASEKAYYA	958
889	GIWVGFKITDPBGYATLGNLELVEEGPLSGDALERLQRBEQWQKIQWTRREERTDRRYMA	948
959	AKQAI DRLFPADYQOQKLSNGVHMSDMLAAQHLVQSIPIYVNDALPEIPGANYTSFTELTN	1018
949	SKQAVDLRYADYQOQLNPDEIITDITAAQBLIQSIPIYVNEPMFPEIPGANYTKFTELTD	1008
1019	RLQQAOWLXLNRNAIPNGDPFNGLSDDNATSDVNVQQLSDTSLVLVIPNMSQVSQQFTVQ	1078
1009	RLQQAOWLITDQRNAIPNGDFPNGLSNNNATPGVEVQQLNHTSVLVIPNWDQVSQQFTVQ	1068
1079	PNRYVLVRITARKEGVGDGYIIRDQGANQTETLTENICDDDTGVLSDAQTSYITKTVFT	1138
1069	PNQRYVLVRITARKEGVNGYVIRDDGNQSETLTFESADYDTNGVYNDQGYITKTVFTI	1128
1139	PSTSQVWIDMSETGVRNIESVELVLBBE	1167
1129	PYTDQMIIEISTEGTGYIBSVELVIVDVE	1157

RESULT 5

RESOUR 5	
AAW06418	
ID	AAW06418 standard; protein; 1157 AA.
XX	
AC	AAW06418;
XX	
AC	
XX	
DT	16-OCT-2003 (revised)
DT	28-JAN-1997 (first entry)
XX	
XX	
DE	Antiscarab pest toxin 50C(a).
XX	
XX	
KW	Toxin 50C(a); scarab pest; toxin; insect; scarabaeidae; pest control;
KW	larval stage insect; grain; tuberous crop; white grub; chafer grub;
KW	cyclocephala; popillia.
XX	
XX	
QS	Bacillus thuringiensis; strain kumamotoensis.
XX	
FN	US5554534-A.
XX	
PD	10-SEP-1996.
XX	
PF	30-SEP-1994; 94US-00315468.
XX	

Qy	479	LSRGYSRHLSHITSYFSKNASSPARVGNLVPFAWTRSDADVTWTVYSDKITQIPVVKAH	538
Db	478	VAESYSRHLSHITSHSFSRGS--AYTSGSPFVFWTHTSDADLNTTYSDKITQIPAVKGD	535
Qy	539	TLVSGTTVIKGPGTGCNLIKRTSSGGLAVTSSVSKPSLSORVYRARIYASTNLRFLVFT	598
Db	536	MLYLGSSVWGPGTGGDILKRTNPSILGTPAVTVNGSLSQRYVRIRYASTTDFB-PTL	594
Qy	599	ISGTRIYSINNVKNTWKGGDILTFNTPDLATIGTAPTFSNYSDSLTVGADSPASGGEVYVD	658
Db	595	YLGDTIBKRNPKNTMDNGASLTFTFPASFTIDPQFRETQDKILLSMGDFSSGQRYIID	654
Qy	659	KPELIPVNATPEAREBIDLVAKKA VNGLPFTSKD DALQTSVTDYOVNOAANLVECLSDSLYP	718
Db	655	RIEPIPWDETYEARBQDLBAKKA VNALPTWTKOGLRPGVTDYEVNOAANLVECLSDSLYP	714
Qy	719	NEKRLMDVAKBAKRLVQARNLLQDTGFNRINGENGMTSGTGEVABGDVLFKDRSLRLT	778
Db	715	NEKRLLPDAVREAKRLSGARNLLQDPDFQBRINGENGMASGTBIEVGDDAVFKGRYLRLP	774
Qy	779	SAREIDTETPTLYLQOIDRSLKDPYTKYKLGPIGSSQDLLEIKLIRHRANOTIVGNVDPN	838
Db	775	GAREIDTETPTLYQYKVEBGLKPYTKYRURGVFGSSQGLEIYTIKHTNRIVKNNVDD	834
Qy	839	LLPDLVPVNSCGGIDRCSQOQYVDANLALENNGENGNMSSDSHAFSPHIDTGEIDLNENT	898
Db	835	LLPDVSPVNSDGSINRCSEQKYNSRL-----EGENRSGDAHEFLPIDIGELDYNEA	888
Qy	899	GIWVVFKEPTTNGVATIGNLELVEBGLSGETLRAQQOQOQODKVARKGAASEKAYYA	958
Db	889	GIWYGFKITDPBGVATIGNLELVEBGLSGDALERLQREQQWKIQMTRERREBETRYMA	948
Qy	959	AKQAIIDLFPADYQDKLNSGVEMSDMLAAOMLVQSIPIVYNDALPEIPGANYTSFTBLTN	1018
Db	949	SKQAVDRLYADYQDOQLNPVBEITDLFRAQDLIGSIPIVYNEMPEIPGANYTKFTELTD	1008
Qy	1019	RLQQAANLYDLRNAIPNGDFPRLNGUSDNNATSDVNVQOISDTSLVLVIPNWSQVSOQFTVQ	1078
Db	1009	RLQQAANLYDORNAIPNGDFPRLNGLSNNVATPGVEVQOQINHTSVLVIIPNWDQVSOQFTVQ	1068
Qy	1079	PNVYVLVRVATKRGVGDGYVIRIDGANOETLLTFNICTDDDTGVLSDAQTSYTKTKTBEFT	1138
Db	1069	PNQRYVLVRVATKRGVGVNGYVSRDGGNQSELTTFPSASDYDVTNGVYNDQGYTKTKTFTFI	1128
Qy	1139	PSTEQVWIDMSFTEGVFNIESVLVLVLEE	1167
Db	1129	PYTDOMWIESETEGTFVIBSVELIIVDVE	1157

RESULT 6

AB027343	
AA027343	
ID	AA027343 standard; protein; 1157 AA.
XX	
XX	
XX	AA027343;
XX	
DT	20-MAY-1998 (first entry)
XX	
XX	B.thuringiensis PS50C Coleopteran-active toxin.
DE	
XX	
KW	B.t.PS50C; NRRL B-18746; pMYC1638; Colorado potato beetle; ss.
XX	
OS	Bacillus thuringiensis.
XX	
PN	CA2059242-A.
XX	
XX	17-JUL-1992.
PD	
XX	
PP	13-JAN-1992; 92CA-02059242.
XX	
PR	16-JAN-1991; 91US-00642112.
PR	02-JAN-1992; 92US-00812180.
XX	
PA	(MYCO) MYCOGEN CORP.

Db 949 SKQAVDLRYADYQDLQLPNDVEITDLTAQOLIQSIPIVYVNEKPEIPGMNTKFTELTD 1008
 Qy 1019 RLQQAANLYDLRNALPNCDFRNGLSDNWATSDVNVQQLSDTSVLVPIPNNSQVSQOFTVQ 1078
 Db 1009 RLQQAANLYDQRNALPNCDFRNGLSNWNATPGVEVQQINHTSVLVI PNWDQVSQOFTVQ 1068
 Qy 1079 PNRYVLVLTARKEGVGGYVIRGANGQTETLTENICDDDTGVLSDADQTSYITKTVFT 1138
 Db 1069 PNQRYVLVLTARKEGVGGYVIRGANGQSELTFTSASDYDTNGVYVNTQVYITKTVFT 1128
 Qy 1139 PSTQVWTDMSSETGVFNIESVELVLEE 1167
 Db 1129 PYTDQMIEISSETGTFYIESVELVDVE 1157

RESULT 8

AAR32354
 ID AAR32354 standard; protein; 1149 AA.

XX AAR32354;

DT 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 14-JUN-1993 (first entry)

XX Coleoptera toxin from B.t. serovar japonensis variety Buihui.

XX Delta-endotoxin; larvae; Anomala cuprea; spore; crystal; Bacillus;

KW thuringiensis; virus; insect; FERM BP-3465.

XX Bacillus thuringiensis; serovar japonensis variety Buihui.

XX WO9303154-A1.

XX 18-FEB-1993.

XX 31-JUL-1992; 92WO-US006404.

XX 02-AUG-1991; 91JP-00193810.

PR 23-JUL-1992; 92US-00915203.

XX (MYCO) MYCOGEN CORP.

PA (KUBI) KUBOTA CORP.

XX Ohba M, Iwahana H, Sato R, Suzuki N, Ogiwara K, Sakanaka K;

PI Hori H, Asano S, Kawabuchi T;

XX WPI; 1993-076511/09.

DR N-PSDB; AAQ36866.

XX New strain of Bacillus thuringiensis serovar japonensis - producing toxin active against coleoptera larvae.

XX Claim 5; Page 28; 48pp; English.

XX The protein sequence is that of a toxin active against Coleoptera that is produced from a pure culture of Bacillus thuringiensis serovar japonensis variety Buihui (FERM BP-3465). The toxin is a delta-endotoxin which has a mol. wt. of ca. 130 kD. It is useful for control of coleoptera larvae e.g. it is effective against Anomala cuprea but has little effect on Lepidoptera. The toxin can be used as B.t. spores or crystals, as opt. treated cells (B.t. or transformed microorganisms) or it is expressed by plants. (Updated on 25-MAR-2003 to correct FN field.) (Updated on 24-OCT-2003 to standardise OS field)

XX Sequence 1149 AA;

XX Query Match 60.6%; Score 3662.5; DB 2; Length 1149;

XX Best Local Similarity 63.8%; Pred. No. 1.1e-232; Indels 61; Gaps 21;

XX Matches 752; Conservative 109; Mismatches 257;

Qy 1 MSPNNQNEYIILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSEGENPELFGNP 60

Db 1 MSPNNQNEYIILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSEGENPELFGNP 60
 Qy 61 EFTFSS-STVGTGIVGQVLCALGVPPAGQIASPVPIVQGLVPSSTSVVEMIMKQVE 119
 Db 61 GTFISAQAVAGTGDIVSTIISGLGIPVLGEVFSILGSLIGLLWFSNNENWQIPANRVE 120
 Qy 120 DLIDQKIDTSVRKALAGLQGLDGLDYYQSKLKNWLENRNDTRARSVVVYQYIALELDF 179
 Db 121 ELIDQKILDSVRSRAIADLANSRJAEVYYQNALERDKRNPHSTRSAALVKERFGNAEAL 180
 Qy 180 VAKIPSPAI SQGVPLLSVYAAQANLHLLLDASIFGAENGFTFGEISTFYDQVTRTA 239
 Db 181 RTNMGSPSQNTYETPELLPTTAAQASLHLLVRDVOIQYKKGWGPONDIDLPHYKQVSYTA 240
 Qy 240 QYSDYCVKWNVTGLDKLGTNAASVLYKHOFREMTLLVDLVALFPNYDTTYTETTA 299
 Db 241 RYSDHCQVWYNAGNLKGTAKQWYDNRFRNNVWVLDLVALFPNYDARIYPLETNA 300
 Qy 300 QLTREVTDPPIVFNRETSGGFCRRWSLN-----SDI-----SPSEVESAVIRSPHLFDIL 349
 Db 301 ELTREIPTDPV-----GSVVTGQSTLISWYDMIPAALPSFSTLEN-LLRKPOFFILL 352
 Qy 350 SEIFFTTRAGLPLNNT-EYLEVNVGHSIKYKNTNASSALERNYGTITSNKIKYDLANK 408
 Db 353 QEIRMYTS---FRQNGTIEYNYMGQRILTSLYIYGSSF--NKYSGVLGAEDIIIPVQGN 407
 Qy 409 DIFQVRSGLADLANVYAGVGPVYASFTLADKNTGSGVGGFTYKPHPTMVCQNYNT 468
 Db 408 DIYRV--VMTYIGRYTNSLLGNVPVTF-YFNSNTQK-----TYSKP-----KQFAGGIKT 454
 Qy 469 ID---EIPPENEPLSRGSHRLSHITSYSPSKNASSPARYGNLVPFAMTHRSADVTNYY 525
 Db 455 IDSGBELTYEN---YQSYSHRVSYITSPEIKSTGTV--LGVPPIFGMTHSSASRNNPIY 509
 Qy 526 SDKITQIPVVAHTLVSGTTVIKGPQ-PTGNNILKRTSSGPLAY---TSVSVKSPLSQRY 581
 Db 510 ATKISQIPINKASRTSGGAVNWFQGLYNGGVPWKLSGSGSQVINLRVAIDAKG-ASQRY 568
 Qy 582 RARIYASTTNLRLFTIIS-----GTRIYSINNVKTNKGGDITFNFDLATIGTA-F 633
 Db 569 RIRIYASDRAGK--TISRSPENPATYSIAIYNTMTWNTASLTYSYFAESGPINL 626
 Qy 634 TFSNYSDSLTVGADSPASGGVYVDKFLIPVNAFTFAEEDLDVAKAVNGLFTSKDAL 693
 Db 627 GISGSSRTFDISITKEGAANLYIDRIBFIPVNTLFAEEDLDVAKAVNGLFTSKDAL 686
 Qy 694 QTSVTDYQVNAANLVECLSDLELYNEKMLWDAVKEAKLYOARNLLQDTGPNRNGEN 753
 Db 687 QTSVTDYQVNAANLIECLSDLELYNEKMLWDAVKEAKLYOARNLLQDTGPNRNGEN 746
 Qy 754 GWTGSGTEVAEGDVLKDRSLRLTSAREIDTETPTLYQQIDSLKLPYTRYKLGPI 813
 Db 747 GWTGSGTEVEVEGDVLKDRSLRLTSAREIDTETPTLYQQIDSLKLPYTRYKLGPI 806
 Qy 814 GSSQDLBKILIRHRANQIVKXNVPDNLPLVFNVSCGIDRCSEQQYVDANLALNNGEN 873
 Db 807 GSSQDLBKILIRHRANQIVKXNVPDNLPLVFNVSCGVDRCSEQQYVDANLALNNGEN 866
 Qy 874 GNMSSDSHAFPHIDTGBIDLNENTGIIWVFKIPTNGVATIGNLEVEEGPLSGETLER 933
 Db 867 GNMSSDSHAFPHIDTGBIDLNENTGIIWVFKIPTNGVATIGNLEVEEGPLSGETLER 926
 Qy 934 AQOQEQWQDMARKRGASKEYAQAQIADRLPADYQDQKLSGVSEKSMALAAQLVQS 993
 Db 927 AQOQEQWQDMARKRGASKEYAQAQIADRLPADYQDQKLSGVSEKSMALAAQLVQS 986
 Qy 994 IPIYVNDALPEIPGMNTYSPTELNRLOQANWLYDLRINAI PNGDFRNGLSQNNATSDVNV 1053
 Db 987 IPIYVNDALPEIPGMNTYSPTELNRLOQANWLYDLRINAI PNGDFRNGLSQNNATSDVNV 1046
 Qy 1054 QQLSDTSVLVPIPNNSQVSQOFTVQPNRYVLVLTARKEGVGGYVIRGANGQSELTFTSASDYDTNGVYVNTQVYITKTVFT 1113

Db 1047 QQLSDTSVLVIPNNSQVSOQFTVPQNYRVLRTARKEGVDGYVIIRDGANQTETLTP 1106
 Qy 1114 NICDDDTGVLSDADTSYITKTVEFTSTEQVWIDMSSE 1152
 Db 1107 NICDDDTGVLSDADTSYITKTVEFTSTEQVWIDMSSE 1145

RESULT 9
 AAR51692
 ID AAR51692 standard; protein; 1149 AA.
 XX AAR51692;
 AC AAR51692;
 XX
 DT 16-OCT-2003 (revised)
 DT 04-NOV-1994 (first entry)
 XX
 DE B.thuringiensis serovar Japonensis insecticidal protein.
 XX
 XX insecticidal protein; Coleoptera larvae; Buibui fungus.
 XX
 OS Bacillus thuringiensis; (serovar Japonensis).
 OS strain Buibui).
 XX
 FH Key Location/Qualifiers
 FT 751..766
 FT /note= "amino acids 751-766 are not given in the
 FT specification and so have been decoded from AAQ58975"
 XX
 XX JP06065292-A.
 PN
 XX
 PD 08-MAR-1994.
 XX
 XX 11-AUG-1992; 92JP-00213886.
 XX
 PR 11-AUG-1992; 92JP-00213886.
 XX
 XX (KUBI) KUBOTA CORP.
 PA
 XX
 DR WPI: 1994-121220/15.
 DR N-PSDB; AAQ58975.
 DR
 FT Insecticidal protein and DNA from Bacillus thuringiensis serovar
 FT Japonensis strain Buibui - useful in insecticides against Coleoptera
 PT insects.
 PT
 XX
 PS Claim 1; Page 9-13; 18pp; Japanese.
 XX
 CC This insecticidal protein has activity against Coleopteran insect larvae
 CC and has been isolated from Bacillus thuringiensis serovar japonensis
 CC strain Buibui. (Updated on 16-OCT-2003 to standardise OS field)
 XX
 XX Sequence 1149 AA;
 SQ

Query Match 60.6%; Score 3662.5; DB 2; Length 1149;
 Best Local Similarity 63.8%; Pred. No. 1.1e-232;
 Matches 752; Conservative 109; Mismatches 257; Indels 61; Gaps 21;

Qy 1 MSPNNQVEYILDASSSTVSQNSVRYPLANDQTTTLONNNYKDYLRMSEGEHPFLGXP 60
 Db 1 MSPNNQVEYIIDALSPTSQNSIRYPLANDQTTTLONNNYKDYLRMSEGEHPFLGXP 60
 Qy 61 ETPISS-STVOTGIGVGVGLGALGVPPFAGQIASFYSGVQLWPSSTVSVMIMKQVE 119
 Db 61 GTPISAQDVGIGIDIVSTIGLIPVLGVFSIILGSLGLLPNNENWQIFPWRVE 120
 Qy 120 DLIDQKIDTSVRKTLAQLGGLDGVYQKSLKNLNNRNDTRARSVVVTVYALELDF 179
 Db 121 ELIDQKIDTSVRKTLAQLGGLDGVYQKSLKNLNNRNDTRARSVVVTVYALELDF 180
 Qy 180 VAKIPSPALSGQVPLLSVQAQANLHLLLRDASIPGAEWFTPGESTFTYDQVTRTA 239
 Db 181 RTNMGSGFSQNTYETPLLPYTAQAASLHLLVRDQVQYKGEWGPQNDIDLFTYKEQVSYTA 240

Qy 240 QYSDYCVKMYNTGDLKLGKTNAAASWLKYHOFREMTLLVLDLVALPFPNYOTRTYPIETTA 299
 Db 241 RYSDHCQVQWYAGLNLKLGKTNAAASWLKYHOFREMTLLVLDLVALPFPNYOTRTYPIETTA 300
 Qy 300 QLTREYVTPDPIVFNRETSGPCRRWSLN-----SDI-----SSEVESAVIRSPHLDL 349
 Db 301 ELTREIFTDPV-----GSYVTGQSTLISWYDMIPALPSPSTLEN-LURKEDFPFTLL 352
 Qy 350 SEIFBYTTRAGLPLNNT-EYLEYVWVGHISIKYKNTNASSALLERNYGTITSNKIKYDILANK 408
 Db 353 QELRWYTS---FRQNGTIEYNTWGGQRLTSLIYIGSSP--NKYSGVLAGAEDIIPVQCN 407
 Qy 409 DIPQVRLGADLANIYAQVYVPYASFTLLDKNTSGSGVGGPYFKPHTTHQVCTQYNT 468
 Db 408 DIYRV--VMYTIIGRYTNSLGLVNPVTP-YFSNNYTK-----TYSK-----KOPAGGIKT 454
 Qy 469 ID---EIPPEBPLSRGYSRSLSHITSYSPSKNASPARYGNLPVPATWTHRSADVTWY 525
 Db 455 IDSGEELTYEN---YQSYSHRVSIIITSFEIKSTGTV--LGVPIPGWTHSSASRNFIY 509
 Qy 526 SDKITQIPVVKATLVSGTIVIKPGP-FTGGINLKRTSSGLAY---TSVSVKSPLSQRY 581
 Db 510 ATKISOIPINKASRTSGGAVWNFQEGLYNGPVMKLGSGSQVINLRVATDAKG-ASQRY 568
 Qy 582 RARIRVASTNLRFLVTVIS-----GTRIYSINVNKTMKGDDLTFTNFDLATIGTA-P 633
 Db 569 RIRIRVADRAGKF--TISRSPENPATYSASLAYTNMTNSTNASLTYSFAYASGPNL 626
 Qy 634 TFSNYSDSLTVGADSPASGGEVYVDKPELIPVATPEAEEDLVAKKAVNGLFTSKKDAL 693
 Db 627 GISGSSRTFDISITKEAGAAANLYIDRIEFPVNTLFEAEEDLVAKKAVNGLFTSKKDAL 686
 Qy 694 QTSVTDYQVNOAANLVECLSDLEYFNEKRLMDAVKEAKLVQARNLLQDTGFNRINGEN 753
 Db 687 QTSVTDYQVNOAANLVECLSDLEYFNEKRLMDAVKEAKLVQARNLLQDTGFNRINGEN 746
 Qy 754 GMTGSTGIEVAGDVLFKDRLTSAREIDTETPTLYLQOIDEISLLKPYTRYKLGPI 813
 Db 747 GMTGSTGIEVAGDVLFKDRLTSAREIDTETPTLYLQOIDEISLLKPYTRYKLGPI 806
 Qy 814 GSSQDLLEIKLIRHRANOIVKQVNDLLPVLVNSCGGIDRCSSQOYVDANLLENNGEN 873
 Db 807 GSSQDLLEIKLIRHRANOIVKQVNDLLPVLVNSCGGIDRCSSQOYVDANLLENNGEN 866
 Qy 874 GNMSSDSHAFSHIDTGEIDLANTGTIWWFKIPTTNGVATLGNLELVBEGLSGETLER 933
 Db 867 GNMSSDSHAFSHIDTGEIDLANTGTIWWFKIPTTNGVATLGNLELVBEGLSGETLER 926
 Qy 934 AQOQEQWQDKMARKGASEKAYAAKQAIIDRLPADYQDOKLNSGVMSDLAAQNLVQS 993
 Db 927 AQOQEQWQDKMARKGASEKAYAAKQAIIDRLPADYQDOKLNSGVMSDLAAQNLVQS 986
 Qy 994 IPVYNDALPEIRPMNYTSPTELNRLLQQAANLYDLRNLNIPNGDFRNLGSDMNATSDVW 1053
 Db 987 IPVYNDALPEIRPMNYTSPTELNRLLQQAANLYDLRNLNIPNGDFRNLGSDMNATSDVW 1046
 Qy 1054 QQLSDTSVLVIPNNSQVSOQFTVPQNYRVLRTARKEGVDGYVIIRDGANQTETLTP 1113
 Db 1047 QQLSDTSVLVIPNNSQVSOQFTVPQNYRVLRTARKEGVDGYVIIRDGANQTETLTP 1106
 Qy 1114 NICDDDTGVLSDADTSYITKTVEFTSTEQVWIDMSSE 1152
 Db 1107 NICDDDTGVLSDADTSYITKTVEFTSTEQVWIDMSSE 1145

RESULT 10
 AAR59256
 ID AAR59256 standard; protein; 1210 AA.
 XX AAR59256;
 AC AAR59256;
 XX
 DT 07-OCT-2002 (first entry)
 XX

Bacillus thuringiensis Cry1218-2 protein sequence.

Pesticidal; spraying; dusting; broadcating; seed coating; insect pest;
Colorado potato beetle; western corn rootworm; southern corn rootworm;
insect target range; endotoxin; Cry1218.

Bacillus thuringiensis..

WO200234774-A2.

02-MAY-2002.

24-OCT-2001; 2001MO-US045468.

24-OCT-2000; 2000US-0242838P.

23-OCT-2001; 2001US-00032717.

(DUPO) DU PONT DE NEMOURS & CO E I.

Abad AR, Duck NB, Feng X, Flannagan RD, Kahn TW, Sims LE;

WPI; 2002-519178/55.

N-PSDB; ABK87235.

New isolated pesticidal polypeptide useful for impacting insect pest e.g.
Colorado potato beetle.

Claim 4; Page 103-105; 176pp; English.

The present invention relates to a new pesticidal polypeptide. The invention is useful for impacting an insect pest by applying the molecules of the invention to the environment of the insect pest by spraying, dusting, broadcating, or seed coating, where the insect pest is selected from Colorado potato beetle, western corn rootworm or southern corn rootworm. The invention is also useful for increasing insect target range and for producing transgenic microorganisms and plants that express the pesticidal polypeptide. The invention is also useful for producing transformed plants and in transforming any organism to produce the pesticidal polypeptide of the invention. The present amino acid sequence represents a *Bacillus thuringiensis* wild-type Cry1218 endotoxin protein

Sequence 1210 AA;

Query Match 54.0%; Score 3263.5; DB 5; Length 1210;

Best Local Similarity 53.4%; Pred. No. 2.6e-206;

Matches 667; Conservative 178; Mismatches 284; Indels 119; Gaps 21;

Qy 1 MSPNNQNEYELDASSSTVSNDNRYPLANDQTTTLQNNYKDYLRMSGENPELFCNP 60

Db 1 MSPNNQNEYELDASSSTVSNDNRYPLANDQTTTLQNNYKDYLRMSGENPELFCNP 60

Qy 61 ETPIS-SSTVGTGIGVGVGALGVPPAGQIASFVIGQLWPSSTVSVMIMKQVE 119

Db 61 EVLVSGQDAKAAIDVGLKLSGLGVFPVGVIVSLYQLDILWPSGQKQWELFMEQVE 120

Qy 120 DLIDQKITDSVRKTAAGLQGLDGLDYQKSLRNLENRNDTRARSVVVYQYIALSLDF 179

Db 121 ELINQKIAEYARNKALSELEGNNYQLYLTALSEWKENPNSREALDVRNRPEILDLP 180

Qy 180 VAKIPSPAIQGVPLSLVYAAQANLHLLLRDASIFGAEWGFTPGELSTFYDQVTRTA 239

Db 181 TQYMFSPRVTFVFPFLTVYTTQANLHLLLRDASIFGEWGWSTTTNNYIDRQMKLTA 240

Qy 240 QYSDCYKRVNTGDKLKGTAASWLYKHQFRREMTLLVLDVALFQNYDTRTYPIETTA 299

Db 241 EYDHCWVYETGLAKLGTSAKQWVDYNQRRMTLTVLDVALFQNYDTRTYPIETTA 300

Qy 300 QLTREVTYDPIVFNRETSGGFCRRWSLNSDISFSEVSASVIRSPHLEILSEIFYTTRA 359

Db 301 QLTREVTYDPIGAVNVSSIG---SW-YDKAPSGVIESVIRPVRVDYITGLVYITQSR 356

Qy 360 GLPLNTEYLEYVWGHISIKYKNTNASSALERNYGTITS-NKIKYDLANKDIFQVRSIGA 418

Db 357 S--ISSARYIRHWAGHOISYHRIFSDNIIKQMYGTNQLHSTSTPFDFTNYDIYKTLSDA 414
Qy 419 DLAN-----YVAQVGVYASFTLLDK--NTGSSGVGGFTYSKPHTTMQVCTONTNID- 470
Db 415 VLLDIVFGTYIYFFGMEFEFFMVLNQLNTRK-----TLKNPYSK 456
Qy 471 -----EIPPE--NEPLSRGYSHRLSHITSYFSKVNASSPARYGNLPVPAWTHRS 517
Db 457 DIIAGTRDSELELPPETSQPNVESYSHRLCHITSIPATGSTT-----GLVPVFSWTHRS 511
Qy 518 ADVNTVYSDKITQIPVYKAHTL---VSG---TTVIGPGFTGNNILKRTSSGPLAYTSV 571
Db 512 ADLNAVHSDKITQIPVYKVSIDLAPSTGTPNNVTWSPGFTGGIIRKVRNG-VIIISHM 570
Qy 572 SVK-SPLSORYRARIYASTTNLRFLVTTISIRYISINVKTNKNGDGLTFTNTDLAGTIG 630
Db 571 RVKISDINKYSKRIYASANTFEPINPSENVKS-HAQTMRGHALTYKNFYATL- 628
Qy 631 TAFTFSNYSDLTVQADSPAS---GGEVYVDKFLIPVNAITFAEEDLDVAKKAVNGLPT 687
Db 629 PPIKFTTTEPPTLGLAIFEAEDFLGIBAYIDRIEPIPVDETYEAEQDLAAKAVNALPT 688
Qy 688 SKKDALQTSYDYQVNOQANLVECLSDLEYPNKGLWDKAEKRLVQARNLLQDTGPN 747
Db 689 NTKDGLPGVYDEVNOQANLVECLSDLLYPNEKRLFLDAVREAKRLSEARNLLQDDPQ 748
Qy 748 RINGENGWGTSGIEVAEGDVLKDRSLRTSAREIDTETPTLYLYQDIDSLKLPYTRY 807
Db 749 EINGENGWGTSGIEVEGDLAFGRVLRPLPGAREIDTETPTLYLYQKVEGVLKPYTRY 808
Qy 808 KLKGFICSSQDLKILRHRAQIVKQVPPDNLDPVLPMNSCGGIDRCSEQQYVDANLAL 867
Db 809 RLKGFVCSQGLAIFTIRHQTNRIVKQVPPDNLDPVSPVNSDGSINRCSEKQVNSRLV 868
Qy 868 EINGENGWGTSGIEVAEGDVLKDRSLRTSAREIDTETPTLYLYQDIDSLKLPYTRY 927
Db 869 ENR-----SGEAHEFSIPIDTGEIDYNENAGIWWGFKITDPEGYATLGNLELVEEGPLS 922
Qy 928 GETLERQQOQOQWODKARKGASEKAYVAAKQAIIDRLPADYQDQKLSNGVMSDMLAA 987
Db 923 GDALERLQREEQKQKIQNTRRBETDRRYKASQAVDLVADYQDQQLNPDVEITDLTAA 982
Qy 988 QNLVQSIPYVNDALPEIPGMYTSTFELTNRLQQAANLVDLRNAINGDFRNGLSDNA 1047
Db 983 QDLIQSIPYVYVNEFPPEIPGMYTSTFELTNRLQQAANLVDLRNAINGDFRNGLSDNA 1042
Qy 1048 TSDVNVQOOLSQTSVLTVPKNSQVSOQTPQPNRYVLYRTARKEGVDGVYIIRDGANQ 1107
Db 1043 TPGVEVQOIHNTSVLVIPIPNDEQVSOQTPQPNRYVLYRTARKEGVDGVYIIRDGANQ 1102
Qy 1108 TETLTFCNICDDDT-GVLSA----- 1125
Db 1103 TETLTFSASDYDNGMYNTQVNTNGYNTNAYNTQASSTNGYNNANNMYTQASNTNGYN 1162
Qy 1126 -----DQSYIYNTYFTSTEQWIDMSETEGVFNIESVELVLEBE 1167
Db 1163 TNSVYNDQTYITNTVTFTFYTDQWMIEMSETEGTFYIESVELIVDVE 1210

RESULT 11

AAU99255

ID AAU99255 standard; protein; 1206 AA.

XX AAU99255;

XX 07-OCT-2002 (first entry)

XX *Bacillus thuringiensis* Cry1218-1 protein sequence.

XX Pesticidal; spraying; dusting; broadcating; seed coating; insect pest;
XX Colorado potato beetle; western corn rootworm; southern corn rootworm;
XX insect target range; endotoxin; Cry1218.

XX 10-SEP-1996.
 XX 30-SEP-1994; 94US-00315468.
 XX 16-DEC-1991; 91US-00808316.
 XX 30-JAN-1992; 92US-00828430.
 XX 01-FEB-1993; 93US-00014941.
 XX (MYCO) MYCOGEN CORP.
 XX Foncerrada L, Narva KE, Michaela TE;
 XX WPI; 1996-424659/42.
 XX N-PSDB; AAT43221.
 XX New nucleic acid encoding B.thuringiensis toxin active against scarab(s)
 PT - also related toxin and transformed microbes, effective against adult
 PT pests and their larvae.
 XX Claim 2; Col 29-36; 24pp; English.
 XX AAW06417-W06419 represent toxins that are active against scarab pests.
 CC The DNA encoding this sequence was isolated from the Bacillus
 CC thuringiensis strain kumamotoensis. Insects in the family Scarabaeidae
 CC constitute a serious pest control problem, especially when destructive
 CC larval stage insects infest high value turf found in golf courses,
 CC playing fields and lawns. The larvae of many species also attack grains,
 CC tuberous crops, and ornamentals. The larvae are known as white grubs, or
 CC chaffer grubs, and can be found in decaying organic matter, or in the soil
 CC where they consume plant roots. In Europe and the U.S. populations of
 CC these larvae and adults have developed resistance to chemical
 CC insecticides such as the organochlorines and DDT. These toxin sequences,
 CC and intact cells that are capable of expressing these proteins, can be
 CC used to control many pests of the family scarabaeidae, such as species of
 CC Cyclocephala, and Popillia. The toxins are active against larvae (present
 CC in soil) and against adults. (Updated on 16-OCT-2003 to standardise OS
 CC field)
 XX Sequence 1169 AA;
 Query Match 52.2%; Score 3152; DB 2; Length 1169;
 Best Local Similarity 54.0%; Pred. No. 5.7e-199;
 Matches 653; Conservative 179; Mismatches 294; Indels 84; Gaps 25;
 Qy 1 MSPNNQNEYILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSEGENPELFGNP 60
 Db 1 MSPNNQNEYILDATPSTSVSDNSNRYPFANEPTNALQNMDDYLYKMSAGNVSEYPCGP 60
 Qy 61 ETPIS-SSTVQTGIGVQVGLGALGVPPAGQIASPYFVQGLWPSSTVSWEMIMQVRE 119
 Db 61 EVFLSEQDAVRAAIDIVGKLTGLGVPPFVGVISLYTLQILDILWPSKQSQWEIPMEQVE 120
 Qy 120 DLIDQKITDVRKTAGLQGLDGLGVYQKSLKNLENRNDTRARSVVVTOYIALELDF 179
 Db 121 ELINQKIAEYARNKALSEGLGNNYQYLTALEEWKENPNSGRALRDVRRRFEILDSIF 180
 Qy 180 VAKIPSPAISQGEVPLLSVYAAANLHLLLRDASIFGAEMGFTFGEISTFYDQVTRTA 239
 Db 181 TQYMPSPRVTVFVPLTVYTTWAANLHLLLRDASIFGEWGLSTSTNNYNNQMKLTA 240
 Qy 240 QYSDYCVKQVNTGLDKLGTWAAWKLKTHQFRREMTLLVLDLWALFPNYDTRTPPIETA 299
 Db 241 EYSDHCWKVYETGLAKLGSAAKQWIDYNQFRREMTLLVLDLWALFNSYDTRTPPLATTA 300
 Qy 300 QLTREYVTDPIVFNRETSQGCRRWSLNSDIVSPSEVASVIRSPLHLDILSEIRPYTTTRA 359
 Db 301 QLTREYVTDPLGADVDPNIG---SW-YDKAFSPSEIEKAAIRPPIVFDITVGLTVYTKR 356
 Qy 360 GLPLNTEYLYWVGHISIKYKNTNASSALERNYGT-----ITSNKIKYDILANKDIFQVR 414
 Db 357 S--FTSDRYMYWAGHQISYKHIGTSSFTQWYGTGNQLQSTN---FDTNYVDYKTL 410

Qy 415 SLGADLAN-----YAAQVGVPIASFTLLDKNTGSGSVGGPTYSKPHTMQVCTQNYNTI 469
 Db 411 SNGAVLLDIVPGYTYTTPFGMPETEFPMVQNLNTRKT--LTY-XP-ASKDIIIDTROSE 466
 Qy 470 DEIPPE--NEPLSRGYSHRLSHITSYSPSKNASSPARYGNLIPVFAWTHRSADVNTVYSD 527
 Db 467 LELPETSQPNYESYSHRLCHIT-PIYSSSTST-----YVPVFSWTHRSADLTWTKSG 520
 Qy 528 KITQIPVVKAHTLVSGTGVINGPGFTGNNILKRTSSGGLPLAYTSVSVKGLPSLORYARIRY 587
 Db 521 EITQIPGKSTIGRNTYIIKRGYTGGLVALDR--IGSCFQMIIPESQRFIRIRY 578
 Qy 588 AS--TTNLRLEP-VTISGTRIYSINVKI-MKGD-DLTFNTFDLATICTAFTFSNYSDSL 642
 Db 579 ASNETSYISLYGLNOSGT-----LKFNTQYSKRNENDLYND-----FYIITPRVI 625
 Qy 643 TVGADS-----PASGEVYV-DKPELIPVNATFEAREEDLDVAKKAVNGLFTSKDAL 693
 Db 626 SVNASSNIQRLSIGIQTNWNLPILDRIEPIPVDETYEARTELEAAKKAVALPTTKDGL 685
 Qy 694 QTSVTDYQVNOAANLVECLSDLELYPNEKGMWDVAKEAKRLVQARNLLQDTPGNRINGEN 753
 Db 686 QPGVTDYEVNOAANLVECLSDLLYPNEKRLLFDAVREAKRLSEARNLLQDPPDQFINGEN 745
 Qy 754 GWTGSGILEVAEGDVLKORSILRTSAREIDTETPTLYQQIDESLLKPYTRYKLKGP 813
 Db 746 GMTASTGIEVIEGDAVFKGRYLRPGAREIDTETPTLYQKVBGVGLKPYTRYLRGV 805
 Qy 814 GSSQDLKELRHRANQIVKVNPNLLPDVLPVNSCGGIDRCSEQYQYVDANLALNNGEN 873
 Db 806 GSSQGLEIYTHQTNRIKVNVPDILLDPVPPVNDGRINRCSEQYKYNLSRLEVENR--- 862
 Qy 874 GMSDSDSHAPFHDITGSDIDANENGIVWPKIPTNGVATIGNLLEVEEGLSGTLER 933
 Db 863 ---SGEAHEPSIPIDTGEIDYNENAGIWWGPKITDPEGVATIGNLLEVEEGLSGDALER 919
 Qy 934 AQQCEQQQDMARKGSEKAYAAKQAIDELADYDOKLNSGVEMSDMLAQNLAOS 993
 Db 920 LQKBEQWKIQMTRRREEDRRYMAKQAVRLYADYDQQLNPNVEITDLTAAQDLIOS 979
 Qy 994 IPYVYNALPEIPGNMYTSTFELTNLQAAWNLVLRNAIPNGDFRNGLSDNATSDVNV 1053
 Db 980 IPYVYNEMPEIPGNMYTKFELTDLQAAWGLYDQRNAIPNGDYRNLSNNTTSGVNV 1039
 Qy 1054 QQLSDTSVLVTPNNSQVSOQTPQPNTRYLVRTARKEGVDGTVIIRDGANQTHETLTP 1113
 Db 1040 QQINHTSVLVIPNWEQVSKFTVQPNQRYLVRTARKEGVNGTVSIRDOGNGQSETLTF 1099
 Qy 1114 NICDDDTGVLSDOTS---YIT-----KTVEFTPSTEQVWIDMSETEGVNI 1157
 Db 1100 SASDYDNGMYDTQASNTNGYNTNSVYMIKPAISRKTVDIVSYNQMWIEISETGTFYI 1159
 Qy 1158 ESVELVLEEE 1167
 Db 1160 ESVELIVDVE 1169
 RESULT 13
 AAW84581
 ID AAW84581 standard; protein; 1157 AA.
 XX AAW84581;
 XX 07-JUN-1999 (first entry)
 XX Amino acid sequence of Bacillus thuringiensis Cry9C mutant 15.
 XX Cry9C; crystal protein; toxic; mutant; insecticide; insect;
 XX transgenic plant; transgenic seed; lepidoptera.
 XX Bacillus thuringiensis.
 XX Synthetic.

PD 07-JAN-1999.
XX
PF 25-JUN-1998; 98WO-BP004033.
XX
PR 27-JUN-1997; 97US-00884389.
XX
XX (PLBZ) PLANT GENETIC SYSTEMS NV.
PA
PI De Roock S, Van Rie J;
XX WPI; 1999-105666/09.
XX
XX New *Bacillus thuringiensis* Cry9C crystal proteins with improved toxicity
PT - useful for producing transgenic plants in controlling pest insects,
PT especially *Ostrinia nubilalis*, *Heliothis virescens* and *Diatraea*
PT *grandiosella*.
XX
XX Example; Page; 40pp; English.
XX
CC This is the amino acid sequence of a Cry9C (crystal) mutant protein which
CC is toxic to an insect species. It is modified in the method of the
CC invention in an attempt to improve its toxicity. The improved proteins
CC are useful for producing transgenic plants and seeds in controlling
CC insects feeding on a plant. The transgenic plants are also useful for
CC producing more transgenic plants with similar characteristics, or for
CC introducing the improved cry9C gene (including fragments) into the same
CC plant variety or related plant species. The Cry9C protein can be obtained
CC from cultured cells of the transgenic plants and used in insecticide
CC compositions against lepidopteran insects, especially *Ostrinia nubilalis*,
CC *Heliothis virescens* and *Diatraea grandiosella*. The Cry9C gene can be
CC expressed in combination with another insect control protein (e.g.
CC Another *B. thuringiensis* derived crystal protein, especially CryIAb- or
CC CryIIb-type protein), and used to transform *B. thuringiensis*, which
CC produces other insecticidal toxins. Such strains are useful against a
CC variety of insect pests, or insects where insect resistance development
CC is prevented or delayed. (Note: this sequence is not given in the
CC specification, but is generated using the information provided by the
CC inventors.)
XX
SQ Sequence 1157 AA;

Query Match 51.8%; Score 3128; DB 2; Length 1157;
Best Local Similarity 52.4%; Pred. No. 2.2e-197;
Matches 620; Conservative 195; Mismatches 326; Indels 42; Gaps 14;

QY 1 MSPNQNEVEILDASSSTSVNSVRYPLANDQTTTLQNNYKDYLRMSGEENPELFGNP 60
DB 1 MNRNNQNEVEIIDAFCGCPSSDDVRYPLASDPNALQNNYKDYLRMSGEENPELFGNP 60
QY 61 ETPFIS-STVQTGIGVQVLCALGVPPAGQIASPVSTVQGLWPSSTVSWEMIMKQVE 119
DB 61 SLSISGRDAVQATLVGRIKALGVPPSGQIVSPYQPLNTLWPNVDTAIWEAFMRQVE 120
QY 120 DLIDQKITSVRKTAGLQIGDGLDYQKSLKWLNRNDTRARSVVVQYIALELDP 179
DB 121 ELVNOQITEFARNQALRQGLGDSFNVYQSLQNLADNRNTRNLVRAQFIALDLDP 180
QY 180 VAKIPSAISQGEVPLLSVYAAALHLLLRDASIFGAEMGFTGCEISTFYDROVTRTA 239
DB 181 VNAIPLPAVNGQVPLLSVYAAALHLLLRDASIFGAEMGFTGCEISTFYDROVTRTA 240
QY 240 QYSDYCVKMYNTGLDKLTNAASMLKYHQPREMTLLVLDLVALFPNVDYTRPIETTA 299
DB 241 KYTNYCETWYNTGLDLRGNTESMLRYHQPREMTLLVLDLVALFPNVDYTRPIETTA 300
QY 300 QLTREVYTPPIVNPRTSGCPERRSLNSDISFSESVAIRSPHLPILSIEFYTTTA 359
DB 301 QLTREVYTPPIVNPRTSGCPERRSLNSDISFSESVAIRSPHLPILSIEFYTTTA 359
QY 360 GLPLNNTLEYVWGHISIKYKNTNASSALERNYGTITSNKIKY---YDLANKDIFQVRS 416
DB 360 -FPV-SSNFMVWGHISIKYKNTNASSALERNYGTITSNKIKY---YDLANKDIFQVRS 416
DB 360 -FPV-SSNFMVWGHISIKYKNTNASSALERNYGTITSNKIKY---YDLANKDIFQVRS 416

QY 417 GADLANYAQTGVPYASFTL--LDKNTSGSGVGFTYSKPHHTMQCTQNTYNTIDBIPP 474
DB 414 AVDFRSALAGIYGVNRASFPVPGELFNGTTSFANGG-----CRDLTYDNDLPP 461
QY 475 ENEPLSRGYS-HRLSHITSYSFSKN-ASSPARVGNLPVPAWTHRSADVNTVYSDKITQI 532
DB 462 DE---STGSSTHLSHVTFPSQTNQAGSIANGSVPTVYVTRDVLANTIIIPNRIQL 518
QY 533 FVVKARTLVSGTVIINGPQFTGNIILKRTSSGLAYTSVSKSPLSQRYDARIYASTTN 592
DB 519 PLVKASAPVSGITVLKPGFTGGILLRRTTNGTEGLRTVNSPLAQOQVRLVRPASTGN 578
QY 593 LRLFTVISTRIYSINNVNTMKGDDLTFTNP---DLATIG---TAFPTSNSDSLTVGA 646
DB 579 FSRVLRGGVSGIDVRLGSTMNRGQELTYESFTTGTGPNPPTFTTQAQEIITVNA 638
QY 647 DFPASGSEVYVDKPELIPVNATFEAREEDLVAKKAVNGLFTSKDKALQTSVTDYQVNOAA 706
DB 639 EGVSTGGTYIIDRIELIVPNPAREAREEDLEAKKAVASLFTTRDGLQNVNTDIYQVDOAA 698
QY 707 NLVCECLSDLYPNEKRMMLMDAVKAEKRLVQARNLLQDTGPNRING--ENGWTCSTGIEVA 764
DB 699 NLVCSLSDQYCHDKMLLEAVRAKLSRERNLLQDPDFTNTINSTEENGKASNGVTIS 758
QY 765 EGDVLFKORSRLTTSAREIDTETYPYLYQQIDESLLKPYTRYKLGKFGTSSQDLKLI 824
DB 759 EGGPFKGRALQALASAR----ENYPTTYQKVDASVLKPYTRYRLDGFVKSSQDLKLI 814
QY 825 RHRANQIVKQVDPNLLPVLVPSNCGGIDRCSEQQVVDANLALENNGENGNSDSHAPS 884
DB 815 HHKRVHLVKNVDPNLSVDYTSVSGSGINRCDBQHVQMDAHHFMDCEAAQTHEPS 874
QY 885 PHIDTGEIDLNTGIVWVFKIPTTNGYATLGNLELVEEPLSGETLERAQQOQEQWODK 944
DB 875 SYINTGDLNASVDQGIWVVLKRVRTDGYATLGNLELVEEPLSGESLERBQRDNKQNAE 934
QY 945 MARKGASEKATYAAKQALIDRLPADYQDQKLSGVEMSDMLAQNLVQSPYVYNDALPE 1004
DB 935 LGRKEAREDRVYLAQAQAINHLFVDYQDQQLNPEIGLAINESNLVBSISGVYSDTLQ 994
QY 1005 IFGMNTSFTELTNRLQQAANLYDLRNALPNCDFRNLGSDWNATSDVNVQQLSDTSVLVI 1064
DB 995 IFGINYEIYELSDRLQQAASYLYTSRNAVQNGDFNSGLSWNTTNDASVQDQGNMFLVL 1054
QY 1065 PNWNSQVSCQFTVQPNRYRYLVRTAKRGVGDGYVIRDGANQTTTLTNI CDDDTGVL 1124
DB 1055 SHWDAQVSCQLRVNPNCKYLVRTARKVGGGEGYVIRDGANQTTTLTNI CDDDTGVL 1114
QY 1125 ADQTSYITKTVEFTSTEQVWIDMSETEGVNIESVELVEEE 1167
DB 1115 VNDNSYITEEVVFPYPTKMMVEVSESGSFYIDSIBFIETQ 1157
RESULT 15
AAW84571
ID AAW84571 standard; protein; 1157 AA.
XX
AC AAW84571;
XX
DT 07-JUN-1999 (first entry)
XX
DE Amino acid sequence of *Bacillus thuringiensis* Cry9C mutant 5.
XX
XX Cry9C; crystal protein; toxic; mutant; insecticide; insect;
KW transgenic plant; transgenic seed; lepidoptera.
XX
XX *Bacillus thuringiensis*.
OS Synthetic.
XX
XX WO990407-A2.
XX
PD 07-JAN-1999.
XX

PF 25-JUN-1998; 98WO-BP004033.
 XX PR 27-JUN-1997; 97US-00884389.
 XX PA (PLBZ) PLANT GENETIC SYSTEMS NV.
 XX PI De Roock S, Van Rie J;
 XX DR MPI; 1999-105666/09.
 XX PT New Bacillus thuringiensis Cry9C crystal proteins with improved toxicity
 PT - useful for producing transgenic plants in controlling pest insects,
 PT especially *Ostrinia nubilalis*, *Heliothis virescens* and *Diatraea*
 PT *grandiosella*.
 XX PS Example; Page; 40pp; English.
 XX CC This is the amino acid sequence of a Cry9C (crystal) mutant protein which
 CC is toxic to an insect species. It is modified in the method of the
 CC invention in an attempt to improve its toxicity. The improved proteins
 CC are useful for producing transgenic plants and seeds in controlling
 CC insects feeding on a plant. The transgenic plants are also useful for
 CC producing more transgenic plants with similar characteristics, or for
 CC introducing the improved cry9C gene (including fragments) into the same
 CC plant variety or related plant species. The Cry9C protein can be obtained
 CC from cultured cells of the transgenic plants and used in insecticide
 CC compositions against lepidopteran insects, especially *Ostrinia nubilalis*,
 CC *Heliothis virescens* and *Diatraea grandiosella*. The cry9C gene can be
 CC expressed in combination with another insect control protein (e.g.,
 CC Another *B.thuringiensis* derived crystal protein, especially CryIAb- or
 CC CryIIb-type protein), and used to transform *B.thuringiensis*, which
 CC produces other insecticidal toxins. Such strains are useful against a
 CC variety of insect pests, or insects where insect resistance development
 CC is prevented or delayed. (Note: this sequence is not given in the
 CC specification, but is generated using the information provided by the
 CC inventors.)
 XX SQ Sequence 1157 AA;

Query Match 51.78; Score 3126; DB 2; Length 1157;
 Best Local Similarity 52.38; Pred. No. 2.9e-157;
 Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;

QY 1 MSPNNQMEYELDASSSTVSNSVRYPLANDQTTLLQNNYKDYLRMSGENSELPFGNP 60
 DB 1 MNRNNQMEYELIDAPHCQCFSDDDVRYPLASDPNAALQNNYKDYLRQWDEYDTSYNP 60
 QY 61 ETPISS-STVQTGIGIVGVGLGALGVPPAGQIASFYSPVIGQLWPSSTVSWEIMKQVE 119
 DB 61 SLGISGRDAVQTALTVVGRILGALGVPPSGQIVSFYQFLNLTLPVNDTAIWEAFMRQVE 120
 QY 120 DLIDQKITDSVRKTALAGLGGLGVYQKSLKNWLENDRTRARSVVVTVQYIALSLDF 179
 DB 121 ELVNNQITFARNQALRQLGLGDSFNVYQKSLKNWLENDRTRARSVVVTVQYIALSLDF 180
 QY 180 VAKIPSPAISSQGVPLLSVYAQAANLHLLLRDASIFCAEWGFTPGBISTFYDRQVTRTA 239
 DB 181 VNAIPLFANNVQGVPLLSVYAQAANLHLLLRDASIFGEGWFTQGHISTFYDRQVTRTA 240
 QY 240 QYSDYCVKWTNTGLDKLKGTTNAAWKLKQFRREMTLLVLDLVALFPNYTRTYPIETTA 299
 DB 241 KTYNYCETWNTGLDRGTNTESWLRHYQFRREMTLLVLDLVALFPNYTRTYPIETTA 300
 QY 300 QLTREVTYTDIVNRETSFGFCRRWSLNSDISSEVESAVIRSPHLEDSIEFYETTRA 359
 DB 301 QLTREVTYTDIVNPPANAGLCRWGNTNPNTSELENAPIRPHLPDLRLNSLTSSNR- 359
 QY 360 GLPLNNTLEYLYVGHGKYGKVTNASSALERNYGTITSNKIKY---YDLANKOIFQVRS 416
 DB 360 -FPV-SNFMDYWSGHTLRSLYNDLSAVQEDSYGLITTTTATINPGVDGTR----LEST 413
 QY 417 GADLANYAAYGVYGVYASFTL--LDKNVSGSGVGFTYKPHFTTMQVCTQNYNTIDRIPP 474

Search completed: June 21, 2004, 13:37:24

Job time : 73 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:39:11 ; Search time 53 Seconds
(without alignments)
6216.227 Million cell updates/sec

Title: US-10-089-678-1
Perfect score: 6044
Sequence: 1 MSPNNQNEYILDASSSTVS.....MSTEGVNIERSVELVLEE 1167

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1163542 seqs, 282313646 residues

Total number of hits satisfying chosen parameters: 1163542

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6044	100.0	1167	12	US-10-089-678-1
2	3263.5	54.0	1210	13	US-10-032-717-4
3	3263.5	54.0	1210	14	US-10-414-637-4
4	3263.5	54.0	1210	16	US-10-606-320-4
5	3246.5	53.7	1206	13	US-10-032-717-2
6	3246.5	53.7	1206	14	US-10-414-637-2
7	3246.5	53.7	1206	16	US-10-606-320-2
8	3100.5	51.3	1156	14	US-10-099-285-72
9	3100.5	51.3	1156	15	US-10-428-961-28
10	2884.5	47.7	1152	15	US-10-614-524-6
11	2825.5	46.7	1150	14	US-10-099-285-74
12	2710	44.8	1134	14	US-10-099-285-76
13	2519.5	41.7	1207	10	US-09-988-462-7
14	2457	40.7	1227	15	US-10-428-961-63
15	2454.5	40.6	1186	9	US-09-826-660-23

16	2449.5	40.5	1228	15	US-10-428-961-38	Sequence 38, Appl
17	2446.5	40.5	1228	15	US-10-614-524-2	Sequence 2, Appl
18	2230.5	36.9	1177	9	US-09-873-873-10	Sequence 10, Appl
19	2230.5	36.9	1177	9	US-09-873-873-12	Sequence 12, Appl
20	2230.5	36.9	1177	9	US-09-873-873-14	Sequence 14, Appl
21	2230.5	36.9	1177	10	US-09-997-914-10	Sequence 10, Appl
22	2230.5	36.9	1177	10	US-09-997-914-12	Sequence 12, Appl
23	2230.5	36.9	1177	10	US-09-997-914-14	Sequence 14, Appl
24	2230.5	36.9	1177	14	US-10-365-645-10	Sequence 10, Appl
25	2230.5	36.9	1177	14	US-10-365-645-12	Sequence 12, Appl
26	2230.5	36.9	1177	14	US-10-365-645-14	Sequence 14, Appl
27	2230.5	36.9	1177	16	US-10-672-163-10	Sequence 10, Appl
28	2230.5	36.9	1177	16	US-10-672-163-12	Sequence 12, Appl
29	2230.5	36.9	1177	16	US-10-672-163-14	Sequence 14, Appl
30	2229.5	36.9	1177	9	US-09-873-873-26	Sequence 26, Appl
31	2229.5	36.9	1177	10	US-09-997-914-26	Sequence 26, Appl
32	2229.5	36.9	1177	14	US-10-365-645-26	Sequence 26, Appl
33	2229.5	36.9	1177	16	US-10-672-163-26	Sequence 26, Appl
34	2227.5	36.9	1163	9	US-09-756-526A-2	Sequence 2, Appl
35	2227.5	36.9	1163	14	US-10-345-020-2	Sequence 2, Appl
36	2227.5	36.9	1163	14	US-10-342-821-2	Sequence 2, Appl
37	2225.5	36.8	1177	9	US-09-873-873-28	Sequence 28, Appl
38	2225.5	36.8	1177	10	US-09-997-914-28	Sequence 28, Appl
39	2225.5	36.8	1177	14	US-10-365-645-28	Sequence 28, Appl
40	2225.5	36.8	1177	16	US-10-672-163-28	Sequence 28, Appl
41	2215.5	36.7	1177	9	US-09-873-873-34	Sequence 34, Appl
42	2215.5	36.7	1177	14	US-10-365-645-34	Sequence 34, Appl
43	2191	36.3	1156	14	US-10-099-285-70	Sequence 70, Appl
44	2154.5	35.6	1174	9	US-09-826-660-6	Sequence 6, Appl
45	2154.5	35.6	1174	12	US-09-837-961-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1

US-10-089-678-1
; Sequence 1, Application US/10089678
; Publication No. US20030017967A1
; GENERAL INFORMATION:
; APPLICANT: ASANO, Shinichiro
; TITLE OF INVENTION: PROTEIN HAVING INSECTICIDAL ACTIVITY, DNA ENCODING THE PROTEIN, AND METHOD
; FILE REFERENCES: Q68821
; CURRENT APPLICATION NUMBER: US/10/089,678
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: JP 2000-236140
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: PCT/JP01/06660
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1167
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-089-678-1

Query Match	100.0%;	Score 6044;	DB 12;	Length 1167;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1167;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	MSPNNQNEYILDASSSTVS	SDNSVRYPLANDQTTLQNNYKDYLRMS	EGENPELFGNP 60
Db	1	MSPNNQNEYILDASSSTVS	SDNSVRYPLANDQTTLQNNYKDYLRMS	EGENPELFGNP 60
Qy	61	ETPISSTVQTGIVGQVILGALGV	PAGQIASFVSPVIGOLWPSSTVS	VWEMIMKQVED 120
Db	61	ETPISSTVQTGIVGQVILGALGV	PAGQIASFVSPVIGOLWPSSTVS	VWEMIMKQVED 120
Qy	121	LIDQKITSVRKKTALAGLGLG	LDVYQKSLKNLENRNDTRARS	VVVVTVYIALELDFV 180
Db	121	LIDQKITSVRKKTALAGLGLG	LDVYQKSLKNLENRNDTRARS	VVVVTVYIALELDFV 180


```
Db 629 PPIKFTTTPITLGAIFEAEDFLGIEAYIDRIEPIVDTEYAEODLEAAKAVNALFT 688
Qy 688 SKDALQTSVTDYQVNOAANLVECLSDLELYPNKRMWDAVKAEXLQVARNLLQDTGFN 747
Db 689 NTKDGLRPGVTDVEVNOAANLVECLSDLELYPNKRMWDAVKAEXLQVARNLLQDPDFQ 748
Qy 748 RINGENGWTSCTGIEVAGDVLKDRSLRLTSAREIDTETPTLYLQOIDESELLKPYTRY 807
Db 749 EINGENGWTSCTGIEVAGDVLKDRSLRLTSAREIDTETPTLYLQOIDESELLKPYTRY 808
Qy 808 KLKGFICSSODLEIKLTHRANOIYKVNPNLLPDLVLPVNSCGGIDRCSQOQVVDANLAL 867
Db 809 RLKGFVSSGQLEIFTHRQNRIVKVNPDLLPDLVSPVNSDGSINRCSQKYNSELEV 868
Qy 868 ENNGENGMSSDSHAFSHDTCEDIDNENGTIWWVFKIPTTNGYATLGNLELVEGPLS 927
Db 869 ENR-----SGEAHEFSIPIDTGEIDYENAGIWWGFKITDPEGYATLGNLELVEGPLS 922
Qy 928 GETLERAQOQOQWQDMARKRGASEKAYAAKQADRLPADYQDQKLNSEVMSDLAA 987
Db 923 GDALERLQREEQWQKIQMTRREBETDRRYMAKQAVDRLYADYQDQQLNPDVEITDLTAA 982
Qy 988 QNLVQSIPYVYNDALPEIPGMNTYTSFTELNRLOQAMNLYDLRNIAPNGDFRNGLSNNA 1047
Db 983 QDLIQSIPYVYNEFFPIPGMYTKFTTELDRLOQAMNLYDLRNIAPNGDFRNGLSNNA 1042
Qy 1048 TSDVNVQQLSDTSVLVTPNNNSOVSOQFTVPQNYRYLVRTARKEGVGDGYVIRDGANQ 1107
Db 1043 TPGVEVQOQIHTSVLVPNDWEQVSQFTVPQNYRYLVRTARKEGVGNGVVSIRDGNQ 1102
Qy 1108 TETLTMICDDDT-GVLASA----- 1125
Db 1103 TETLTSASDYDNGMYNTQVSTNGYNTNNAYNTQASTNGYNANNMYNTQASNTNGYN 1162
Qy 1126 -----DOTSYTKTWTZFTSTEQWIDMSTETGCVNIESVELVLEE 1167
Db 1163 TNSVYNDQGTGTYTKTVPFYPTQMWIEMSETEGTFTYIESVELVDOVE 1210

RESULT 3
US-10-414-637-4
; Sequence 4, Application US/10414637
; Publication No. US20030177528A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Feng
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn B. Sims
; TITLE OF INVENTION: Genes Encoding No. US20030177528A1el Proteins With
; TITLE OF INVENTION: Pesticidal Activity Against Coleopterans
; FILE REFERENCE: 35718/237005
; CURRENT APPLICATION NUMBER: US/10/414,637
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: US/10/032,717
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1210
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
; US-10-414-637-4

Query Match 54.0%; Score 3263.5; DB 14; Length 1210;
Best Local Similarity 53.4%; Pred. No. 3e-258;
Matches 667; Conservative 178; Mismatches 284; Indels 119; Gaps 21;

Qy 1 MSPNNQNEYEIIIDATPSTSVNSDNRYPANETPNALQNMDYKDYLOMSAGNASEYPCSP 60
61 ETPIS--SSTVQGTIGIVGOVLGALGVPPFAGQIAFSFYFVGOQLSPSTSVSWEMIMQVE 119
61 EVLVSGQDAKAAAIIDIVGKLLSGLGVPPFVGVPIVSLYTQIIDILWPSQKQWEIFMEQVE 120
120 DLTDOKITDSVRKTALAGLQGLDGLDVYQKSLKNWLENDRTRARSVVVYQYIALELDF 179
121 BLINQKIAEYARWAKLSALEGLGNNYQLYLTALKEEMKENPNSRALLDRVRFRFILDSLP 180
180 VAKIPSAISQGBVPLLSVYAQAANLHLRLDASIFGAEMGFTFGHISTFYDRQVTRTA 239
181 TQMPSEFVTFVPPFLTVYVYQAANLHLRLDASIFGBEMGWTSTTTNNYYDRQMKLTA 240
240 QYSDYCVKNTNTGLDKLKGNAASWKLKHOFREBMTLLVLDLVALFPNYDTRTPIETTA 299
241 EYSDHCVKWYETGLAKLKGTSKQVNDYQFRBEMTLTVLDVVALFPNYDTRTPIETTA 300
300 QLTREVYVTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSPLHLDLSEIFYTTRA 359
301 QLTREVYVTDPLGAVNVSSIG---SW-YDKAPSPGVIESSVIRPHVFDYIYGLTYTQSR 356
360 GLPLANNTEYLEYVWGHISIKYKNTNASSALERNTGTITS-NKIYYIDLANKDIQVRSIGA 418
357 S--ISSARYIRHWAGHQISYHRIEFDNIIKQMYGTINQLHSTSTFDFNTYDIYKLSKDA 414
419 DLAN-----YYAQVYGVYASFTLLDK--NTGSGSVGGFTYSKQHTTMQVCTQNTTID- 470
415 VLDDIVPPGYTYIYFGHPEVEFPMVNLNTRK-----TLKYNPVSK 456
471 -----EIPPE--NEPLSRGYSRLSHSTYSFSPKNASSPARGNLPVPAWTHRS 517
457 DIIAGTRDSELEPPETSQDFNYESYSHRLCHITSIPATGSTT-----GLVPVFSWTHRS 511
518 ADVTNTVYSDKIQTQIPVWKAHTL---VSG---TVIKGPGTGGNMLKRTSSGPLAYTSV 571
512 ADLINAHSKIQTQIPVWYSDLAPSITGGPNNTVSGPGTGGGIIKVIENG-VIISHM 570
572 SVK-SPLSQRYARIRYASTNTNRLFVTISGTRISYINVNKTKNGKDDLTNTFPDLATIG 630
571 RVKISDINKYSNRIRYASANTTEFYINPSEENVKS-HAQTNRGEALTYNKNFYATL- 628
631 TAPTFSNYSLSLTVGADSPAS---CGEVYVDKPELIPVNATFEABEDLDVAKKAVNGLFT 687
629 PPIKFTTTPITLGAIFEAEDFLGIEAYIDRIEPIVDTEYAEODLEAAKAVNALFT 688
688 SKDALQTSVTDYQVNOAANLVECLSDLELYPNKRMWDAVKAEXLQVARNLLQDTGFN 747
689 NTKDGLRPGVTDVEVNOAANLVECLSDLELYPNKRMWDAVKAEXLQVARNLLQDPDFQ 748
748 RINGENGWTSCTGIEVAGDVLKDRSLRLTSAREIDTETPTLYLQOIDESELLKPYTRY 807
749 EINGENGWTSCTGIEVAGDVLKDRSLRLTSAREIDTETPTLYLQOIDESELLKPYTRY 808
808 KLKGFICSSODLEIKLTHRANOIYKVNPNLLPDLVLPVNSCGGIDRCSQOQVVDANLAL 867
809 RLKGFVSSGQLEIFTHRQNRIVKVNPDLLPDLVSPVNSDGSINRCSQKYNSELEV 868
868 ENNGENGMSSDSHAFSHDTCEDIDNENGTIWWVFKIPTTNGYATLGNLELVEGPLS 927
869 ENR-----SGEAHEFSIPIDTGEIDYENAGIWWGFKITDPEGYATLGNLELVEGPLS 922
928 GETLERAQOQOQWQDMARKRGASEKAYAAKQADRLPADYQDQKLNSEVMSDLAA 987
923 GDALERLQREEQWQKIQMTRREBETDRRYMAKQAVDRLYADYQDQQLNPDVEITDLTAA 982
988 QNLVQSIPYVYNDALPEIPGMNTYTSFTELNRLOQAMNLYDLRNIAPNGDFRNGLSNNA 1047
983 QDLIQSIPYVYNEFFPIPGMYTKFTTELDRLOQAMNLYDLRNIAPNGDFRNGLSNNA 1042
1048 TSDVNVQQLSDTSVLVTPNNNSOVSOQFTVPQNYRYLVRTARKEGVGDGYVIRDGANQ 1107
1043 TPGVEVQOQIHTSVLVPNDWEQVSQFTVPQNYRYLVRTARKEGVGNGVVSIRDGNQ 1102
1 MSPNNQNEYEIIIDATPSTSVNSDNRYPANETPNALQNMDYKDYLOMSAGNASEYPCSP 60
```



```
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1206
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-032-717-2

Query Match      53.7%; Score 3246.5; DB 13; Length 1206;
Best Local Similarity 53.4%; Pred. No. 7.5e-257;
Matches 665; Conservative 172; Mismatches 290; Indels 119; Gaps 18;

QY 1 MSPNNQNEVEILDASSSTVSNSVRYPLANDOTTTTLONNNYKDYLRMSGENPFLGPNP 60
DB 1 MSPNNQNEVEILDATSTVSNSNSRYPPANEPTNALQNMWDYKYLKMSAGNASSEYPGSP 60

QY 61 ETPIS-SSTVQTGIGIVGVGLGALGVPFAGQIASFYSPFVQQLWSPSTVSVMIMKQVE 119
DB 61 EVLVSGQDAKAALDIVGLKSLGLGVFPVGPVIVSLYTLQILWPSGEKSWEIFMEQVE 120

QY 120 DLIDQKITDSVRKTALAGLQGLDGVYQKSLKNWLENDRTRARSVVVTOYIALELDP 179
DB 121 ELINQKIABYARNKALSELEGLGNNYQLYLTALKEEENPNRSLRDVNRFRILDSLF 180

QY 180 VAKIPSPALSGQVPLLSVYAOAANLHLLLRDASIFGAEWGPFGISTFYDQVTRTA 239
DB 181 TQMPSPFRVTFVFPFLTVYMAANLHLLLRDASIFGAEWGPFGISTFYDQVTRTA 240

QY 240 QVSDYCVKMYNTGLOKLGTAASWIKYHQPREMILLVLDLVALPNDYTRTYPIETTA 299
DB 241 EYSDHCYKMYETGLAKLGTSARQWYDQYQFREMTLAVLDVVALPNDYTRTYPIETTA 300

QY 300 QLTREYVYTOPIVNRRTSGCFRWSLNSDISFESEVASVIRSPHFLDILSEIEFTTRA 359
DB 301 QLTREYVYTOPLGAVNVSSIG--SW-YDKAPSGVIESSVIRPHFPDYITGLTVYVTSQR 356

QY 360 GLPLNTEYLEYVWGHISIKYKNTNASSALERNYGTITS-NKIYDYLANKDIPQVRSLGA 418
DB 357 S--ISSARIYRWAGHQISVHRVSRGSLNQMYGTQNLHSTSTPFTNVDYIKYLSKDA 414

QY 419 DLAN-----YYAQYGVVYASFTLLQK--NTGSGSGVGGFTYSKPHTMQVCTQNYTID- 470
DB 415 VLIDVIVPGVYTYFFGKPEVEPEPMVQLNNTRK-----TLKYNPVSK 456

QY 471 -----RTPPE--NEPLSRGYSRHLSHITSYSPKSNASSPARYGNLPVATHRS 517
DB 457 DIIATRDSELEPPETSDQPNYESYSHRLCHITSIPATGNIT-----GLVPVFSWTHRS 511

QY 518 ADVNTNYSKTIQIIPVWKAHLVSGTVIKGPGFGGNILKRTSS-----GPLAYTSVSU 573
DB 512 ADLNTIYSDKIYQIIPAVKCDWNLPPVUVVKGFGHTGGDLLQYNRSTGSGVGLFLARYGL 571

QY 574 KSPLSQRYRARIYASTTNRLFTVTISGTRIYSINVKNTMKNKGGDLTPTNFDLATIGTAF 633
DB 572 ALEKAGKYAVLAYATDADIVLHVDA-----QIQPKTNPCGEDILTSKTEKVAADAITTL 626

QY 634 ---TFPNYSLSLTVGADSPAS--GGEVTVDFKELIPVNAIFEAEDLDVAKAINGLPTSK 689
DB 627 NLATDSSALKRNLGHDPNSTLSGIVYVDRIFIPVDTEYAEQDLEAAKCAVNALFTNT 686

QY 690 KDALQTSVTDYQVQNAANLVECLSDLEYPNKCKMLMDAVKAEARLVQARNLLQDTQFNRI 749
DB 687 KGLRGVTDYEVNQAANLVECLSDLLYPNEKRLLDVAVREARLVQARNLLQDPDFQEI 746

QY 750 NGENGTGSGTIEVASEGDLVLFQDRLSLTSARBIDTETPTLYLQQIDESLLKPYTRYKL 809
DB 747 NGENGTASTGIEVIEGDLAFKGRYLRLPGAREIDTETPTLYLQKVEEGLKPYTRYKL 806

QY 810 KGFIGSSQDLIELKIRHRANQIVNVPDNLPLDVLVNSCGGIDRCSEQQYVDANLLEN 869
DB 807 RGPVGSQGLIEITIRHOTNRIYKQNVDPDLLPDVSPVNSDGSINRCSEQYVNSLEVEN 866

QY 870 NGENGNSSDSHAFSPHIDTGEIDMLNENTGIWVFKIPTTNGYATLGNLHLEVEGPILSGE 929
DB 867 R-----SGEAEHESIPIDTGEIDYNNENAGIWWGFKITDPGEYATLGNLHLEVEGPILSGD 920

QY 930 TLERAQOQOQOQWODKMARKKASAKAYAAKQAIIDLPAFYDQOKLNSGVSWSMKLAQON 989
DB 921 ALERLQREQOQWIKIOWTRREEDTDRRYMASQAVDRLYADYDQOQLAPDVEITDLTAAQD 980

QY 990 LVQSIYVYNDALPEIPGMNYTSFTLITNRLQOAWNLIDLNAIPNGDPFRNGLSDWNATS 1049
DB 981 LIQSIYVYNEPPEIPGMNYTKFTLITDLQOAWSLYDQRNAIPNGDPFRNGLSNWNTATP 1040

QY 1050 DVNVQQLSDTSVLVIVPNNSQVSQQFTVQPNRYVLRVARTARKGSGVGVYIIRGAKQTE 1109
DB 1041 GVEVQOINETSVLVIVPNWDEQVSQQFTVQPNRYVLRVARTARKGSGVGVYIIRGAKQTE 1100

QY 1110 TLTFNICDDDT-CVLSA----- 1125
DB 1101 TLTFASDYDTNGMYNTQVSTNGYNTNNAINTQASTNGYNNANNMTQASNTNGNTNTN 1160

QY 1126 ----DOTSYITKTVPTSTPTEQWIDMSSETGVFNIESVELVLEE 1167
DB 1161 SVYNDQGTITKTVPTPYTDQMIEMSETGVFIIESVELIVDVE 1206

RESULT 6
US-10-414-637-2
; Sequence 2, Application US/10414637
; Publication No. US20030177528A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Feng
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn E. Sims
; TITLE OF INVENTION: Genes Encoding No. US20030177528A1el Proteins With
; TITLE OF INVENTION: Pesticidal Activity Against Coleopterans
; FILE REFERENCE: 35718/237005
; CURRENT APPLICATION NUMBER: US/10/414,637
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: US/10/032,717
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1206
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-414-637-2

Query Match      53.7%; Score 3246.5; DB 14; Length 1206;
Best Local Similarity 53.4%; Pred. No. 7.5e-257;
Matches 665; Conservative 172; Mismatches 290; Indels 119; Gaps 18;

QY 1 MSPNNQNEVEILDASSSTVSNSVRYPLANDOTTTTLONNNYKDYLRMSGENPFLGPNP 60
DB 1 MSPNNQNEVEILDATSTVSNSNSRYPPANEPTNALQNMWDYKYLKMSAGNASSEYPGSP 60

QY 61 ETPIS-SSTVQTGIGIVGVGLGALGVPFAGQIASFYSPFVQQLWSPSTVSVMIMKQVE 119
DB 61 EVLVSGQDAKAALDIVGLKSLGLGVFPVGPVIVSLYTLQILWPSGEKSWEIFMEQVE 120

QY 120 DLIDQKITDSVRKTALAGLQGLDGVYQKSLKNWLENDRTRARSVVVTOYIALELDP 179
DB 121 ELINQKIABYARNKALSELEGLGNNYQLYLTALKEEENPNRSLRDVNRFRILDSLF 180

QY 180 VAKIPSPALSGQVPLLSVYAOAANLHLLLRDASIFGAEWGPFGISTFYDQVTRTA 239
DB 181 TQMPSPFRVTFVFPFLTVYMAANLHLLLRDASIFGAEWGPFGISTFYDQVTRTA 240

QY 240 QVSDYCVKMYNTGLOKLGTAASWIKYHQPREMILLVLDLVALPNDYTRTYPIETTA 299
DB 241 EYSDHCYKMYETGLAKLGTSARQWYDQYQFREMTLAVLDVVALPNDYTRTYPIETTA 300

QY 300 QLTREYVYTOPIVNRRTSGCFRWSLNSDISFESEVASVIRSPHFLDILSEIEFTTRA 359
DB 301 QLTREYVYTOPLGAVNVSSIG--SW-YDKAPSGVIESSVIRPHFPDYITGLTVYVTSQR 356

QY 360 GLPLNTEYLEYVWGHISIKYKNTNASSALERNYGTITS-NKIYDYLANKDIPQVRSLGA 418
DB 357 S--ISSARIYRWAGHQISVHRVSRGSLNQMYGTQNLHSTSTPFTNVDYIKYLSKDA 414

QY 419 DLAN-----YYAQYGVVYASFTLLQK--NTGSGSGVGGFTYSKPHTMQVCTQNYTID- 470
DB 415 VLIDVIVPGVYTYFFGKPEVEPEPMVQLNNTRK-----TLKYNPVSK 456

QY 471 -----RTPPE--NEPLSRGYSRHLSHITSYSPKSNASSPARYGNLPVATHRS 517
DB 457 DIIATRDSELEPPETSDQPNYESYSHRLCHITSIPATGNIT-----GLVPVFSWTHRS 511

QY 518 ADVNTNYSKTIQIIPVWKAHLVSGTVIKGPGFGGNILKRTSS-----GPLAYTSVSU 573
DB 512 ADLNTIYSDKIYQIIPAVKCDWNLPPVUVVKGFGHTGGDLLQYNRSTGSGVGLFLARYGL 571

QY 574 KSPLSQRYRARIYASTTNRLFTVTISGTRIYSINVKNTMKNKGGDLTPTNFDLATIGTAF 633
DB 572 ALEKAGKYAVLAYATDADIVLHVDA-----QIQPKTNPCGEDILTSKTEKVAADAITTL 626

QY 634 ---TFPNYSLSLTVGADSPAS--GGEVTVDFKELIPVNAIFEAEDLDVAKAINGLPTSK 689
DB 627 NLATDSSALKRNLGHDPNSTLSGIVYVDRIFIPVDTEYAEQDLEAAKCAVNALFTNT 686

QY 690 KDALQTSVTDYQVQNAANLVECLSDLEYPNKCKMLMDAVKAEARLVQARNLLQDTQFNRI 749
DB 687 KGLRGVTDYEVNQAANLVECLSDLLYPNEKRLLDVAVREARLVQARNLLQDPDFQEI 746

QY 750 NGENGTGSGTIEVASEGDLVLFQDRLSLTSARBIDTETPTLYLQQIDESLLKPYTRYKL 809
DB 747 NGENGTASTGIEVIEGDLAFKGRYLRLPGAREIDTETPTLYLQKVEEGLKPYTRYKL 806

QY 810 KGFIGSSQDLIELKIRHRANQIVNVPDNLPLDVLVNSCGGIDRCSEQQYVDANLLEN 869
DB 807 RGPVGSQGLIEITIRHOTNRIYKQNVDPDLLPDVSPVNSDGSINRCSEQYVNSLEVEN 866
```

```
Qy 240 QYSDYCVKQYNTGLDKLKGNTNAASWLYKHQFRREMTLLVLDLVALFPNYDTRTYPIETTA 299
Db 241 EYSDHCYKQYNTGLDKLKGNTNAASWLYKHQFRREMTLLVLDLVALFPNYDTRTYPIETTA 300
Qy 300 QLTREYVTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSHLELDILSEIIFYTTRA 359
Db 301 QLTREYVTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSHLELDILSEIIFYTTRA 356
Qy 360 GLPLNNTLEYLYVWGHSHIKYKNTNASSALENRYGITS-NKIYYDLANKDIFQVRSIGA 418
Db 357 S--ISSARIYRHAGHQISYHRVSGNSLQOMYGTONLHSTSTPDTFNYDIYKLSKDA 414
Qy 419 DLAN-----YYAQVGYVPYASFTLLDK--NTGSGSVGGFTYSKPHHTMQVCTQNTID- 470
Db 415 VLLDIVVPGYTYIFPGMEVEFFMVNQLNTRK-----TLKYNPVSK 456
Qy 471 -----EIPPE--NEPLSRGYSHRLSHITSYSFSKSNASSPARVGNLVPFAWTHRS 517
Db 457 DIIASTRDSLELPPETSQDNYESYHRLCHITSIPATGNT-----GLVPFVSWTHRS 511
Qy 518 ADVNTVYSDKITQIPVKAHTLVSGTIVIKGPGFTGNILKRTSS-----GPLAYTSVSV 573
Db 512 ADLNTIYSDKITQIPAVKCDNLPFPVVKPGHGTGDLQYNRSTGSGVTLFLARYGL 571
Qy 574 KSPLSQRYRIRVASTNLTNLFVTISGTRIYSINVNKTMKGGDLTNTFDLATIGTAP 633
Db 572 ALEKAGKYRVLRYATDADIVLHVDA-----QIQMPKTMNPGEDLTSKTEFKVADAITL 626
Qy 634 ---TFSNYSDSLTVGADSFAS-GEHYVDYKELIPVNAFPAEDOLDVAKAVNGLFTSK 689
Db 627 NLATDSSLALKHLNGEDPNSLGIYVVDRIEPTVDETAEQOOLEAKKAVNALPTNT 686
Qy 690 KDALQTSVTDYQVNOAANLVECLSDLPNPKRMWDVAKRAKRAVQARNLQDGENRI 749
Db 687 KQGLRPGVTDYEVNOAANLVECLSDLPNPKRMWDVAKRAKRAVQARNLQDGENRI 746
Qy 750 NGENGWTSSTGIEVAEGDVLKFDRLSRLTSAREIDTETPTLYLQOODESLKPYTRYKL 809
Db 747 NGENGWTSSTGIEVEGDALFKGRYLRLPGAREIDTETPTLYLQOODESLKPYTRYKL 806
Qy 810 KGFTGSSQDLKILIRHANOIVKRNPNILPDVLVNSCGGIDRCSQOQVVDANALEN 869
Db 807 RGFVGSQGLIEFTIRHQTNIRVKNVDDLLPVDVSPVNSDGSINRCSQKTVNSKLEVEN 866
Qy 870 NGENWSSDGHAFSFHIDTGEIDLNENTGIWVFKIPTINGYATLGNLLELVESGPLSGE 929
Db 867 R-----SGEAHEFSIPDTGEIDYNENAGIHWGFKITDPGAYATLGNLLELVESGPLSGD 920
Qy 930 TLERAQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQO 1049
Db 921 ALERLQREEQWQKIQWTRRREEDRRYVASKQAVDRLYADYQOQOQOQOQOQOQOQOQOQOQO 980
Qy 990 LVQSIPTYNALPBI PCNNTTSFTLNTLRQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQO 1049
Db 981 LIQSIPTYNALPBI PCNNTTSFTLNTLRQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQO 1040
Qy 1050 DWNQOQSDTSVILVTPNNSVSOQOFTQPNRYVRLVYARKEGVDGCVYIIRDCANOTE 1109
Db 1041 GVEVQOQIHTSVLVIPNWDQVSOQOFTQPNRYVRLVYARKEGVDGCVYIIRDCANOTE 1100
Qy 1110 TLTNICDDDT--GVLSA----- 1125
Db 1101 TLTFASADYDNGMNTQVNTNGYNTNAYNTQASNTNGYNNAYNTQASNTNGYNTN 1160
Qy 1126 ----DQTSYITKVEFTSTQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQO 1167
Db 1161 SVYNDQAGYITKVTIFPYTDQMIEMSEBTCTFYIESVELIVDS 1206
```

RESULT 7

US-10-606-320-2

; Sequence 2, Application US/10606320

```
; Publication No. US20040091505A1
; GENERAL INFORMATION:
; APPLICANT: Andre E. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnall
; APPLICANT: James P.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/263948
; CURRENT APPLICATION NUMBER: US/10/606,320
; CURRENT FILING DATE: 2003-06-25
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1206
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
; US-10-606-320-2
```

Query Match 53.7%; Score 3246.5; DB 16; Length 1206;

Best Local Similarity 53.4%; Pred. No. 7.5e-257; Indels 119; Gaps 18;
Matches 665; Conservative 172; Mismatches 290;

```
Qy 1 MSPNNQNEYETLDASSSTSVSDNSVRYPLANDQTTTQNMNMYDYLRMSBEGENPELFGMP 60
Db 1 MSPNNQNEYETLDASSSTSVSDNSVRYPLANDQTTTQNMNMYDYLRMSBEGENPELFGMP 60
Qy 61 ETPFIS-SSTVGTGIGVQVLGALGVPPAGQIASFYISVQGLMPSSTSVVWEMTKOVE 119
Db 61 ETVLSGQDAKAAIDIVGKLSGLGVFPVGVISVLTQIDILMPSBKSQMEIFMEQVE 120
Qy 120 DLIDQKTDSVRKLTALAGLQGLGDVYQSKLKNLENRNDTRARSVVVYQYIALELDF 179
Db 121 ELINQKLAERYARNKALSELEGNNYQLYLTALBEEENPNGSRALRDVNRFBILDSLF 180
Qy 180 VAKIPSAISQORVPLISVYQAQANLHLLLRDASIPGAEWGFTTPEISTFTDQRTVTA 239
Db 181 TOYMPSPFVTFVFPVFLTVYANAANLHLLLRDASIPGAEWGFTTINNTYDQRMQLTA 240
Qy 240 QYSDYCVKQYNTGLDKLKGNTNAASWLYKHQFRREMTLLVLDLVALFPNYDTRTYPIETTA 299
Db 241 EYSDHCYKQYNTGLDKLKGNTNAASWLYKHQFRREMTLLVLDLVALFPNYDTRTYPIETTA 300
Qy 300 QLTREYVTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSHLELDILSEIIFYTTRA 359
Db 301 QLTREYVTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSHLELDILSEIIFYTTRA 356
Qy 360 GLPLNNTLEYLYVWGHSHIKYKNTNASSALENRYGITS-NKIYYDLANKDIFQVRSIGA 418
Db 357 S--ISSARIYRHAGHQISYHRVSGNSLQOMYGTONLHSTSTPDTFNYDIYKLSKDA 414
Qy 419 DLAN-----YYAQVGYVPYASFTLLDK--NTGSGSVGGFTYSKPHHTMQVCTQNTID- 470
Db 415 VLLDIVVPGYTYIFPGMEVEFFMVNQLNTRK-----TLKYNPVSK 456
Qy 471 -----EIPPE--NEPLSRGYSHRLSHITSYSFSKSNASSPARVGNLVPFAWTHRS 517
Db 457 DIIASTRDSLELPPETSQDNYESYHRLCHITSIPATGNT-----GLVPFVSWTHRS 511
Qy 518 ADVNTVYSDKITQIPVKAHTLVSGTIVIKGPGFTGNILKRTSS-----GPLAYTSVSV 573
Db 512 ADLNTIYSDKITQIPAVKCDNLPFPVVKPGHGTGDLQYNRSTGSGVTLFLARYGL 571
Qy 574 KSPLSQRYRIRVASTNLTNLFVTISGTRIYSINVNKTMKGGDLTNTFDLATIGTAP 633
```

Db 572 ALEKAGYRVRRLRYATDADIVLHVND-----QIQPKTWNPGEDLTKFKVADAJTTL 626
Qy 634 ---TFSNYSLSLTGADSPAS-GGEVYVVDKPELIPVNAATPEABEDLDVAKKAVNGLFTSK 689
Db 627 NLATDSSLAKHNLGEPNSTLSGIYVVDRIEPIPVDETTYEAEQDLAKKAVNALFTNT 686
Qy 690 KDALQTSVTYQVNOAANLECLSDLELYPNKRMMLDAVKEAKLVQARNLLQDTGNRI 749
Db 687 KGLRPGVTDEYVNOAANLECLSDLELYPNKRLLPDAVREAKLSEARNLLQDPQEI 746
Qy 750 NGENGTSGTGIKVAEGDVLFKORSILTSAREIDTETPTLYLQOIDSLLKPYTRYKL 809
Db 747 NGENGTASTGIEVIEGDALFKGYRLPLGAREIDTETPTLYLQYKEGVLKPYTRYL 806
Qy 810 KGFPGSSQDLLEKLRHRAQIVKQVNDLIPVNSCGIDRCSEQQVVDANLALEN 869
Db 807 RGFVSSQGLEIFIRHQTWRIVKQVNDLIPVNSDGSINRCSEQKVNRSLEVEN 866
Qy 870 NGENGNSDSHAPSFHIDTGEIDLNERVTGIWVFKIPTNGYATLGNLELVESPLSGE 929
Db 867 R-----SGEHRFSIPIDTGEIDYNENAGIUVGFKITDPEGYATLGNLELVESPLSGD 920
Qy 930 TLERAQOQEQOQKQWARKGASEKAYYAAQKQIDRLPADYQDQKLSGVNEMSLAAQN 989
Db 921 ALERLQBEQOQKQWARKGASEKAYYAAQKQIDRLPADYQDQKLSGVNEMSLAAQN 980
Qy 990 LVQSIPIVYNDALPEIPIQVNTSFTLTNRLQQAANLYDLRNLIPNGDFRNLGSLDMNATS 1049
Db 981 LIQSIPIVYNDALPEIPIQVNTSFTLTNRLQQAANLYDLRNLIPNGDFRNLGSLDMNATS 1040
Qy 1050 DVNVQQLSDTSVLVPIPNNSQVSOQFTVQPNRYRLVLTARKEGVGDGYVLIIRGANOTE 1109
Db 1041 GVEVQINHTSVLVI PNWDEQVSQFTVQPNRYRLVLTARKEGVGDGYVLIIRGANOTE 1100
Qy 1110 TLFPNICDDDT-GVLSA----- 1125
Db 1101 TLTPSASDYDTNGMYNTQVNTNGYNTNAYNTQASSTNGTNGNANMYNTQASNTNGYNTN 1160
Qy 1126 ---DQTSYTKTVEFTSPTEQWIDMSSETGEVFNIESVELMLVEE 1167
Db 1161 SVYNDQGYTKTVEFTSPTEQWIDMSSETGEVFNIESVELMLVEE 1206

RESULT 8

US-10-099-285-72
Sequence 72, Application US/10099285
Publication No. US20030105319A1

GENERAL INFORMATION:
APPLICANT: Schnepf, H. Ernest

Wicker, Carol
Narva, Kenneth E.
Walz, Michelle
Stockhoff, Brian
Muller-Cohn, Judy

TITLE OF INVENTION: Toxins Active Against Pests
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:

ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/099,285
FILING DATE: 15-Mar-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/002,285
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/886,615
FILING DATE: 1-JUL-1997
APPLICATION NUMBER: US 08/674,002
FILING DATE: 1-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-701C2
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 1156 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-10-099-285-72
Query Match 51.3%; Score 3100.5; DB 14; Length 1156;
Best Local Similarity 52.0%; Pred. No. 6.9e-245;
Matches 615; Conservative 196; Mismatches 330; Indels 41; Gaps 13;
Qy 1 MSPNNQNEYELDASSSTSVSDNSRYPLANDQTTTLQNNMYKYLRMSEGENPELFGNP 60
Db 1 MNRNNQNEYELIDAPHCPCSDDDVRPLASDPNAAQLQNNMYKYLTQWDEDTDSYNP 60
Qy 61 ETPISS-STVQGTGIGIVGVLGALGVPPAGQIASFYFVGLQMPSSSTSVVEMTKQVE 119
Db 61 SLISIGRDAVQTALTVVGRILGALGVPSGQVYFQELLTLPVNDTALWEAFMRQVE 120
Qy 120 DLIDQKLTDSVRKTAGLQGLDGLDYVQKSLQNLNENRNDTRASVVTQYIALELDF 179
Db 121 ELVNNQITTEFARNQALRQGLGDSFNYYQKSLQNLNENRNDTRASVVTQYIALELDF 180
Qy 180 VAKIPSPASIGQKVPVLLSVTAQAANLHLLLRDASIFGAERMGFTFGEISTFVDQVTRTA 239
Db 181 VNAIPLFAVNGQVPLLSVTAQAANLHLLLRDASIFGAERMGFTFGEISTFVDQVTRTA 240
Qy 240 QYSDYCVKQYNTGLDKLGTNAASMLKYHOPRRMTLLVLDLVALFPNYOTRTYDIETTA 299
Db 241 KYTNYCETWNTGLDLRGNTESWLYHQFRMTLLVLDLVALFPNYOTRTYDIETTA 300
Qy 300 QLTREVTDDIVNRETSGGFCRRWSLNSDISFSEVESAVTIRSHPHLPDILSIEPVTTRA 359
Db 301 QLTREVTDDIVNRETSGGFCRRWSLNSDISFSEVESAVTIRSHPHLPDILSIEPVTTRA 359
Qy 360 GLPLNNTVLYRYVGHSHIKYKNTNASSALERNYGTITSNKIKY---VDLANKDIFQVRLS 416
Db 360 -FPV-SSNFDYSGHTLRRLSYLNDQSAVEDSYGLITTRATINPGVDGTR-----IBST 413
Qy 417 GADLANYTAQVYGVYASFTL--LDKNTGSGVSGGFTYKPHHTQVCTQYNTIDRIPP 474
Db 414 AVDFRSALIGYGVNRASFPVPGGLFNGTTSFANGG-----CRDLYDTNDELPP 461
Qy 475 ENBPLSRGYS-HRLSHITSYSPSKN-ASSPARYGNLPVPAWTHRSADVNTVYSDKITQI 532
Db 462 DB---STGSSTHRLSHVTFPFPQTNQAGSTANAGSVPTVYVTRDVLNNTIPNRIYQL 518
Qy 533 PVVKAHTLVSGTIVKGFPGTGNILKRTSGPLATYSVSKSPLSQRYRARIYASTN 592
Db 519 PLVKASAPVSGTIVKGFPGTGNILKRTSGPLATYSVSKSPLSQRYRARIYASTN 578
Qy 593 LRLPVTISGTRIYSINVNKTNKGDDLTFTNFDLATIGT-----AFTPSYSDSLTVGAD 647
Db 579 PSIRILRGNTSIAYQRFSGTMRNGELTYESFTSBEPTNQSDLPFTFTQAGNLTILAE 638
Qy 648 SPASGGEVYVDKPELIPVNAATPEABEDLDVAKKAVNGLFTSKDALQTSVTIDYQVNOAN 707


```

; Mon Jun 21 15:02:27 2004
; us-10-089-678-1.rapb
; Page 9

RESULT 10
US-10-614-524-6
; Sequence 6, Application US/10614524
; Publication No. US20040016020A1
; GENERAL INFORMATION:
; APPLICANT: Arnaut, Greta
; APPLICANT: Boets, Annemie
; APPLICANT: Damme, Nicole
; APPLICANT: Mathieu, Eva
; APPLICANT: Vaneste, Stijn
; APPLICANT: Van Rie, Jeroen
; TITLE OF INVENTION: Insecticidal proteins from Bacillus thuringiensis.
; FILE REFERENCE: NEWTSUS2
; CURRENT APPLICATION NUMBER: US/10/614,524
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: US/09/739,243
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/173387
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-614-524-6

Query Match 47.7%; Score 2884.5; DB 15; Length 1152;
Best Local Similarity 49.1%; Pred. No. 3.8e-227;
Matches 578; Conservative 200; Mismatches 365; Indels 33; Gaps 14;

QY 1 MSPNQNEYILDSASTSVSDNSVRPLANDQTTTLQNNYKDYLSMSGEPELPGNP 60
DB 1 MNRNQNDYEIDNASGCGCASDDVQVPLARDNPVQNNHYKDYLTQYDGYTGSFNP 60
QY 61 ETPISSSTV-QTGIGIVGVGLGALGVPPAGQIASFYSFIVGQLWPSSTVSWVMIMKQVE 119
DB 61 NLSINPRDLVQGINIVGRLLGLFVPPAGQVLTFTFLLNQLKPTNDNAWEAFMAQIE 120
QY 120 DLIDQKITDSVRKALAGLGLDGLDVYQKSLKNMLENNDTRFARSVVVTVYIALELDF 179
DB 121 ELINQISEAVGTAADHLTGLHDNVELYVEALEEWELEPNAART-NLLENRPFTLDSLP 179
QY 180 VAKIPSFAL----SCQEVPLISVYAQAANLHALLARDASIFGAEMGTPGCEISTFYDROV 235
DB 180 TQPMPSFGTGPSONYAVPLLATVYAQAANLHLLLDKABYIGARWGLNQNSPHTROQ 239
QY 236 TRTAQYSDYCVKWTGLDKLQGTNAASWLKYHOFREMTLLVLDLVALPPNYDTRTPI 295
DB 240 ERTQYTNHCVTYNTGLDLRGNTESWLNYHFRFREMTLMAVDLVALFPYTNVQYPN 299
QY 296 ETTPAQLTREYITDPIVFNRETSGGFCRWSLNSDISFSEVASVIRSPHLFDILSRIFY 355
DB 300 GANPQLTREYITDPIVVPNPANQICRKGWNNPNTFSELENAPIRPPLHFLDLNRLTIS 359
QY 356 TTRAGLPLANTEYLETWGHSIKYKVNASSALERNYGTITSNKIKYDLANKDIFQVRS 415
DB 360 RNRYPATPNS-YLDYWSGHTLQSOYANNPTTETSYGQITSN-TRLFNTTN-GANAIDS 416
QY 416 LGADLANYAQVYGVVPYASFTLLDKNTGSGSVGGFTYKSKPHTTMQVCTQNYNTIDEIPE 475
DB 417 RARNFGLYANLYGVSYLNI-----PPTGWSEIT-SAEWT-----CQDLTTTBEPLV 465
QY 476 NEPLSRGYSHLSHTSYFSKNASFP-ARYGNLPVPAWTHRSADVNTVYSKTIQIPV 534
DB 466 NNNP-----NLLSEVTLFLENTTQGGELATVGVFTVVTQDVNDENNIITPNRIQIPV 520
QY 535 VKAHTLVSGTIVIKGPGTCGNILKRTSSGLATVTSVSKSPLSQRVARIIRVASTNLR 594
DB 521 VKAYELSSGATVVKGPGTGGDVIRNTGTGFAIRSVTVGTGFLTQVRIIRVASTIDED 580

595 LFWTISGTRIYSINNVNKMANKGDDLTNTFEDLATIGTAFTPSNYSDSLSLTVGADSPASGGE 654
581 FFWTRGGTTINFRFTTRMNRGQESRYESYKTVFTTTPFNFQTOSQDIIRTSIQGLSGNGE 640
655 VYVDKPELIPVNATFRAEEDLDVAKKAVNGLFTSKDALQTSVTDYQVNOAANLVECLSD 714
641 VYLDRIIIPVNPTRAEEDLEAAKAVASLFTTRDGLQVNVTDYQVQQAANLVSCLSD 700
715 ELYPNEKRLMDVAKKRLVQARNLLQDTCFNRING--ENGWTCSTGIEVAEGDVLFPD 772
701 EQYAHDKMLLEAVRAAKRLSRENTLQDPDFNTINSTEENGWAKASNGVTISEGGPFYKG 760
773 RSLRLSARIDTETPTTYLQQIDESLLKPYTRYKLGKPGICSSODLEIKLIRHANOIV 832
761 RALQLASAR---ENYPTTYQKVDASELKPYTRYRLDGFVKSSQDLIEDIHHKHLV 816
833 KNPVDNLLPDVLPVNSCGGIDRCSEQQYVDANLALENGB-NGNMSSDSHAPSFIHDIGE 891
817 KNVLNLSVDTYDDSCSGINRCBEOQMVNAQLTEHHHPMDCCCAAQTHFESSYIDYGD 876
892 IDLNTGTVWVFKIPPTNGYATLGNLLEVEEGPLSGETTLERAAQOQOQOQOQOQOQOQO 951
877 LNSTVDQGIWVIFKVRTTDCGYATLGNLLEVEVGLLGEPLEREQRENKNAELGRKRAE 936
952 SEKAYAAKQAIIDLRLPADYQOQKLSNGSVESDMLAONLVOSIPYVYNDALPEIPGMVY 1011
937 TDRYQDAKQSIHNLFDYQOQLNPQIGMADIMDAONLVAISIDSVYSDAVLQIPGINYE 996
1012 SFTLNLRLQOANLVDLRNAIPNGDFRNLSDWNTSDVNVQQLSDTSVLVLPNNSQV 1071
997 IYTELNLRLQOASLYLTSRNAVQNGDPNGLSDSNATAGASVQOQDGNTHFLVLSHWAQV 1056
1072 SQOFTVQPNRYVLRVTRARKEGVGDGVYIIRDGANQOTETLTFTNCDDDTGVLSADQTSYI 1131
1057 SQFVQPNCKYVLRVTRARKEGVGDGVYIIRDGANQOTETLTFTNCDDDTGVLSADQTSYI 1131
1132 TKTVEFTPTSEQWIDMSFTGVNIESVELVLEER 1167
1117 TKEVIFYSHTEHVVNEVTEGAFHIDSIEFVTEK 1152

RESULT 11
US-10-099-285-74
; Sequence 74, Application US/10099285
; Publication No. US20030105319A1
; GENERAL INFORMATION:
; APPLICANT: Schnepf, H. Ernest
; Wicker, Carol
; Narva, Kenneth E.
; Walz, Michelle
; Stockhoff, Brian
; Muller-Cohn, Judy
; TITLE OF INVENTION: Toxins Active Against Pests
; NUMBER OF SEQUENCES: 105
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/099,285
; FILING DATE: 15-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/002,285
; FILING DATE: <Unknown>
```

APPLICATION NUMBER: US 08/886,615
FILING DATE: 1-JUL-1997
APPLICATION NUMBER: US 08/674,002
FILING DATE: 1-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-701C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 74:
SEQUENCE CHARACTERISTICS:
LENGTH: 1150 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 74:
US-10-099-285-74

Query Match 46.7%; Score 2825.5; DB 14; Length 1150;
Best Local Similarity 48.6%; Pred. No. 2.6e-222;
Matches 574; Conservative 197; Mismatches 364; Indels 47;

Qy	1	MSPNNQNEBYEILDASSSTSVSDNSVRYPLANDQTTTLQNMNTKOYLRSBGE-----NP	54
Db	1	MNRNPNPEYEIIDAQPYCCPDDDDYFVLASDPNAAQNMNYKEYLQTYDGDYGTGSLINP	60
Qy	55	ELFGNPETFISSSTVQTGIGIVGOVLGALGVPPAGQIASFSYFI VQGLWPSSTVSVWEMI	114
Db	61	NLSINPR-----DVLQCTGINIVGRILGFLGVPPAGQLVTFYFYLNLQNLWPTNDNAWEAP	115
Qy	115	MKQVEDLIDQIKTDSVRKXTALAGLQGLGDLDVYOKSLKNMLLENRNDTRARSVVVTVXIA	174
Db	116	MAQIEELIDQIKSAQVRNALDODL TGLHDYVEEYLAAL EEWLERPWCARA- NLVTRQRPEN	174
Qy	175	LELDQVAKIPSPATIS- GQ8-----VPLISVYQAQANLHLLLLDASIFGAHWGTPCEISIFP	230
Db	175	LHTAPVTMPGFGTQSGQORDAVALLVYQAQANLHLLLLDKDAEYIGARWGLQOQGINLY	234
Qy	231	YDRQVTRTAQTSDYCVKRWYNTGLDKLGTNAASWLKYHQFREMTLLVLDLVALEPNYDT	290
Db	235	PNAQQRTRITYNTHCVETNRYGLEBVRGTNTESMLNYHRPREMTLMAMDVLVLPFFYNV	294
Qy	291	RTYPIETTAQLTREYVTPDPIVFNRETSGFCERWSLNSDI SPSEKVESAVIRSPHLPDILS	350
Db	295	ROYENGANPOLTREITYTPDPIVYNPPANGQICRRWGNPNYTFSELENAPIRPPHFLPERLN	354
Qy	351	EIEFVYTRAGLPLANNTEYLEVYVGHISIKYKQTNASSALERNYGTITSNKIKYYDLANKDI	410
Db	355	RLTISRNRYPAPTNS-FLDYWSGHTLQSQHANNPTTETS YGQITSN-TRLFNTTN-GA	411
Qy	411	FOVRSGLADLANYAQVGVGYASFTLLDKQVCGSGVSGGYTSKXPHITMAQVCTQNYNIID	470
Db	412	RAIDSRARNFGNLYANLYGV--SSLNIPP--TG-----VMS EITNAANTCRODLTTFTE	460
Qy	471	EIPENEPILRGYSHRLSHITSYSPSKNASSP-ARYGMLPFWAMTHRSADVNTNTVYSDKI	529
Db	461	ELPLENNPF----NLLSHVIFLRFNITQGGELATLGFVPTVYVWTRDVDFTNITADRI	515
Qy	530	TOI PVVKAKHTLVSGTIVYIKGCPFTGCNLLKRTSSGPLAYTSVSVKSPISQRYRARIIRVAS	589
Db	516	TQLEPWKASEITGGGTVYVKGPGFTGDDTLRLRTDGGAVGTRANVNAPI TQOYRILRYLAS	575
Qy	590	TTN--LRLPVTISGTRIYSINVKTMNKGDDLT FNWTFDLATIGTAPFTSNYSDSLTUGAD	647
Db	576	TTSFVNVLFWNNSAA--GFTLPSTMAQNGSLITYESFNWTFVTHTRFSQSDTTLRLINP	632
Qy	648	SFASGGEVYVDKFULIPVNATFEAREDLDVAKAVNGLFTSKDQALQTSVTDYOVNQOAN	707
Db	633	PSISGQBYVDKLEIVNPNPTFEAREDEDKAKAVASLFTTRDGLQVNWVYQVDOQAN	692

TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-81000
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 1134 amino acids
TYPE: amino acid
STANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO
US-10-099-285-76

Query Match	44.88;	Score 2710;	DB 14;	Length 1134;
Best Local Similarity	48.16;	Pred. No. 7.9e-213;		
Matches 559;	Conservative 195;	Mismatches 345;	Indels 62;	Gaps 19;
Qy	39	NNMYKDYLRMSBEGNPBL-----FGNPETPISSSSTVQTGIGIVGVQVLGALGVPPAQOIA	92	
Db	4	NPNINBCIPYNCLSNPEVVLGSGRNVRT-----GLQTGIDIVAVVVGALGPGVGGILT	58	
Qy	93	SFYSPIVQQLWPSTSVSWEMIKQVBDLIDOKITDSVRTKALAGLOGLDELGVOKSL	152	
Db	59	GFLSTLFGFLWPSNDQAVWEAFIQMBELIEQRISDQVVRTALDDLTGIONVYNOYLIAL	118	
Qy	153	KNWLENRNDTRARSVVYTOVIALBLDPVAKIPSPAISGO-----EVPLLSVYQAANLHL	207	
Db	119	KWEERENGVRA-NLVLRQREILHALFVSSNPSPG-SGPGSQRFQAQLLVYQAANLHL	176	
Qy	208	LLLRDASIPQARWGFPTGRISTFVDRQV-TRTAQYSDYCVKMYNTGLDKLGTNAAGWLK	266	
Db	177	LLLDAAEKYGARWGLRSQIGNLYFNEFLQTRDRTYTHCVNAYNNGLAGLGTSAESWLK	236	
Qy	267	YHOPRREMTLLDLVALPNYDRTYPIETTAQLTRVEYTDPIVFENRETSGGP-----CR	322	
Db	237	YHOFRRRATLWMDLIALFPYNTRRYPYIAVNPQLTRVEYTDPIGVSPSEBSLPEELRCL	296	
Qy	323	RWSLNSDISFSEVASIRSPHLFDLISEBFTYTRAGLPLNNTEYLEYVWYGHSI--KYK	380	
Db	297	RMQETSAWTFSENENAIISPHLFDTLNNLMIXTGFSPSVHLTN-QLIEGWIGHSVTSLL	355	
Qy	381	NTYASSALERNYGHITSNKIKYYDLANKDIPQVRS---LGADLANYAQVYGVYPASFLL	437	
Db	356	ASGPTVILRLNRYGTSSTIS-IVNYFSPFNDRVYQINTRSHGTGLGFQN--APLEGITRAQF--	410	
Qy	438	LDRKNTGSCSGFTYSKPHITMQVCTQNYNTIDBIP--PENEPLSRGYSHRLSHITSYSP	495	
Db	411	-----YPGGYSYVQURNALTCBQNYNSIDHLPDLFNEPISRSYSHRLSHITSYH	461	
Qy	496	-----SKNASSPARYGNLPVPAWTHRSADVNTVYSDKIQTQIPVWKAHTLVSGTTVKG	549	
Db	462	RVLTIDGINIYS-----GNLPYVWTHRDVDLTNTIATDRITQLPLVKSFBIPAGTIVVRG	517	
Qy	550	PGFTCGNHLKETSGLPLAYTSVSKSPLSORSYRARIYASVTNLRFLPTISGTRIYSINV	609	
Db	518	PGFTGGDILRTGTGTGTRVTRTAPITQBYRIRFPASTNIFIGIRGDRQVYDFP	577	
Qy	610	NKTWKGDDLTFTNTFDLATIGTATFTFSNYSDSLTVGADSPASGGEVYVDFKELIPVNAVF	669	
Db	578	GRTWNRGDELRYESPATREPTTDFNRPQPQELISVFANAFSAGQGVYFDRIEIIIPVNPAP	637	
Qy	670	EABEDLVAKAVNGLFTSKDQALQTSVTDYQVNOQANLVCELSDELYFPNKRMLWDVAK	729	
Db	638	EAKEDLBAAKKAVASLFTTRTDGLQVNVKYQVQQAANLVCSLSDBOGYDKQKMLLEAVR	697	
Qy	730	EAKRLVQARNLLQDTGFNRING--ENGWTSGTGTEVABEGDVLFKDRSLRLTSAREIDTET	787	
Db	698	AAKCLSRERNLLQDPDFWNTINSTENGWKAANGVTISEGGPFYKGRALQIASAR-----EN	753	
Qy	788	YPTYLYQQO1DBSLKPYTRYKLGFGTSSQDLKIKLIRHRANQIVKNVDPNLLDPVLVFN	847	
Db	754	YPTYIYQKVDASELKPYTRYRSDGFKVSSODLEIDLHHRKVLHVKNVDPNLSVSDYTPDD	813	

D	B	813	VESPIGRCGEPRNCA--PHFEMNPDLDCSDCRDGBKCAHHSHHFLSDIDVGCTDLHENLGV	870
Q	Y	901	WVFKIPPTNGVATIGNELVREGPISEHTLERAAQQOQOOQODXVARQRGASBKAYPAAK	960
D	B	871	WVVFKKIKTOBGEHARLGNTLEFEKPLGLGEALSRVRAEKKMRDKSELQLQLETKRVYTEAK	930
Q	Y	961	QAIDRLFPADYQDOKLNSGVMSDMAAQNVLQSIPVVYNDAIPEIPGNMYTSFTBLTNRL	1020
D	B	931	EAVDALFVDSQDRLOADTNIQHHAOKLVIRIREAYLSELFWIPGVNAEIFEELEGHI	990
Q	Y	1021	OQAWNLVDLRNAIPNGDFRAGLSDNWATSDDNVQOLSDTSVLVIPNNMSQVSQQFTVQPN	1080
D	B	991	ITAISLYDARNVVKNGDFNNGLTWNVRGHVDVQCSHRSLDIVPEWEABYSQAVRVCPG	1050
Q	Y	1081	YRVURVTARKSGVGGVYVIRDGANQTETLP-----	1113
D	B	1051	CGYILRVTAYPEGYEGGCVTTHIEINNTDELKFKNRBEERYPTDTGTCDNYTAHQGTAG	1110
Q	Y	1114	-----NICDDTGVLSSAD--QTSY	1130
D	B	1111	CACACNSRNAGYDAYEVDDTASVNVKYTYBEETTVDRDNHCYDRGYVNPVPVPGY	1170
Q	Y	1131	ITKTEFTPSTEQWIDMSETGEVENIESVELVLEE	1167
D	B	1171	VYKELEYPPETVWIBIGETEGKPIVDSVELLMEE	1207
 RESULT 14 US-10-428-961-63 ; Sequence 63, Application US/10428961 ; Publication No. US20030237111A1 ; GENERAL INFORMATION: ; APPLICANT: Baum, James A. ; APPLICANT: Chu, Chih-Rei ; APPLICANT: Donovan, William P. ; APPLICANT: Gilmer, Amy J. ; APPLICANT: Rupar, Mark J. ; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin ; FILE REFERENCE: MECO201--1 ; CURRENT APPLICATION NUMBER: US/10/428,961 ; CURRENT FILING DATE: 2003-05-02 ; PRIOR APPLICATION NUMBER: 09/661,322 ; PRIOR FILING DATE: 2000-09-13 ; PRIOR APPLICATION NUMBER: 60/153,995 ; PRIOR FILING DATE: 1999-09-15 ; NUMBER OF SEQ ID NOS: 63 ; SOFTWARE: PatentIn version 3.2 ; SEQ ID NO 63 ; LENGTH: 1227 ; TYPE: PRF ; ORGANISM: Bacillus thuringiensis US-10-428-961-63				
 Query Match 40.7%; Score 2457; DB 15; Length 1227; Best Local Similarity 42.8%; Pred. No. 5.4e-192; Matches 541; Conservative 197; Mismatches 392; Indels 134; Gaps 20;				
Q	Y	1	MSPNQNEVEILDASSSTSVDNSVRYPLANDQTTLQNMVYKDYLRMSGENPELPCNP	60
D	B	1	LTSNRKENBHIALUSIPAVSNHQMMLSTD-----ARIEDSLCAREG-----NNI	47
Q	Y	61	ETFFISSSTVQTGIGVGOVLGAIVYPFAQIASFYSFIVGOLWPSSTSVSMEMIMQVED	120
D	B	48	DPFVSASTVQGINAGRILGVYPFAQIASFYSLVGLGLWPRGR-DPWEIFLHVHEH	106
Q	Y	121	LIDQKITDSVRKTALAGLGGLEDLDVTVQKSIAKNWLENRDTRARSVVTVQVIALEDFV	180
D	B	107	LIRQOVTEMTROTALARLQGLGNSFRAYQOSLEDWLENRDDARTSFLVTVQVIALEDFL	166
Q	Y	181	AKIPSPAISGSQRPVLLSYVAQAAHLALLLRDASIPGAENGFTPCBISTIFYDVORTAQ	240
D	B	167	NAMPLAFARNOSVPLAMVYAQAAHLALLLRDASLFGSEFGTSGEQIRYYERTKTRRE	226

```
Qy 241 YSDYCVKMYNTGLDKLKGNTNAASMLKXHQFREMFLVLDLVALFPNYDTRTYPIETTAQ 300
Db 227 YSDYCARWYNTGLNLRGTNAESWLRYNQFRDLTLGLVLDLVALFPSTYTRVYPMNTSAQ 286
Qy 301 LTRVYTDPIVFNRETSGPCERHSLNSDISFSVESAVIRSHPLFDILSEIEFTTAAQ 360
Db 287 LTRVYTDPIVFNRETSGPCERHSLNSDISFSVESAVIRSHPLFDILSEIEFTTAAQ 345
Qy 361 LPLANTLEYWVGHSHIKYKNTNASSALERNVGTITSNKIKYDILANKDIFQVRSIGADL 420
Db 346 -RWSTQYMYWVGHSHIRSLTSTWTHGNTNTSINPVLTQFTSRDVRTSPAGIN 404
Qy 421 ANYYAQVGVYPAFTLLDKNTGSGVGGFTYSKPHHTMQVCTQNTYNTIDIPPE--NEP 478
Db 405 ILLTPVGVFWARFNN--RNPLNSLRGSLLYTYGT--GVGTQLFDSSETLPPETTRP 460
Qy 479 LSRGYSRLSHITYSFSKKNASSPARYCN---LVPANTHRSADVNTVYSDKTIQIPV 535
Db 461 NYESYHRLSNIRLIS-----GNTLRAPVYSWTHRSADRTNTISSDITQIPLV 509
Qy 536 KAHTLVSGTIVKPGFTGGNLRKTSGLPLAYTSVSVKSPLSQRYRARIYASTNLR 595
Db 510 KSFNLNGSTSVSGPGTGGDIIRTNVNGSVLSGLNPNNTSLQRYRVRYAASQTWVL 569
Qy 596 FVTISGTRIYSINVNKTKNGDDLTFTNFDLATIGTAFTPSNYSDSLTVGADSPASGGEV 655
Db 570 RVTVGGSTTFDQGPSTMSANESLTSQSFRAEPFVGISASG--SQTAGISINNAGRTF 628
Qy 714 YVDFEELIPVNTATEAEDLDVAKKAVNGLTSKKD--ALQTSVTDYQVQNAANLVECLSD 714
Db 629 HFDKIEFIPITATFEAYDLERAQEAVALFTNTNPRRLKGTVDYHIDEVSNLIVACLSL 688
Qy 715 ELYPNEKMLMDVAKEAKRLVQANLLQDTGFNRIN-----GENG 754
Db 689 EFLDDEKRELLKVKYAKGLSDERNLLQDPNFTSINKQDPNNSNEQSNFTSIHQSEHG 748
Qy 755 WTGSTGIEVAGDVLFKDRSLRLTSARBIDTETPTLYLYQOIDSLLKPYTRYKLKPGTG 814
Db 749 WQGSENITIQENDVFXENYVTLPGT---FNCEPTLYLYQKIGRAELKAVTRYQLSGVTE 805
Qy 815 SSQDLLEIKLIRHRANQIVKVP--DNLLPDVL--PVNSCGGIDRCSGQYVDANLALENN 870
Db 806 DSQDLLEILIRYNAKHETLDVPGTESVWPLSVESPIGRGCEPNRCA--PHFENAPDLDCS 863
Qy 871 GENG--NMSSDSHAPSFDITGEIDLNENTGIWVFKIPTNGYATIGNLHLEVBEGPLSGR 929
Db 864 CRDEKCAHSHHPSLDIVGCDLHENVGVVFKITQEGHARLGNLEFIEKPLIGE 923
Qy 930 TLERAQOQOQOQOQOQMARKGASBKAYAAKQADRLFADYQDQKINSVEMSDMLAQN 989
Db 924 ALSRVKRAEKWRDKREKLQLETKRVYTEAKEAVDALFVDSQYDRLQADTNIGMHAADK 983
Qy 990 LVQSIPIVYNDALPEIICMNYTSFELTNRLQOAWNLVLENAJPNCGDFRNLSDWATS 1049
Db 984 LVHREAYLSLSVIFGVNAEIIPEELREGRITITAISLYDARNVYVNGDFNNGLACWVKG 1043
Qy 1050 DVNVQOOLSLSVLIPIPNNSQVQFTVQPNRYVLRVLTARKEGWDGYVLIIRQANOTE 1109
Db 1044 HVDVQOQSHRSVLVPIPEAEVSQAVRCPGRGILRLVATYKEGVBGGCVTIHBIENND 1103
Qy 1110 TLTFNLCDD-----DTGVLS-----ADQTS----- 1129
Db 1104 ELKPKNCEEBEYPTDGTGTCNDYTAHQCTAVCNSRNAGYEDAVEVDVTTASVNYKPTYEEB 1163
Qy 1130 -----YITKVFETPTEQWIDMSHTEGVNFIESVELV 1163
Db 1164 TYTDVRDNHCEYDRGYVYPPVPPAGYMTKELEYFPEPDKVWIBIGETGKFIIVDSVELL 1223
Qy 1164 LEEB 1167
Db 1224 LMEB 1227
```

```
RESULT 15
US-09-826-660-23
; Sequence 23, Application US/09826660
; Patent No. US20010026940A1
; GENERAL INFORMATION:
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Stelman, Steven J.
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
; FILE REFERENCE: MA-714XC2D1
; CURRENT APPLICATION NUMBER: US/09/826,660
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/178,252
; PRIOR FILING DATE: 1998-10-23
; PRIOR APPLICATION NUMBER: 60/065,215
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/076,445
; PRIOR FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 23
; LENGTH: 1186
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Toxin encoded by synthetic B.t. gene
US-09-826-660-23
```

```
Query Match 40.6%; Score 2454.5; DB 9; Length 1186;
Best Local Similarity 43.1%; Pred. No. 8.3e-192;
Matches 536; Conservative 191; Mismatches 382; Indels 135; Gaps 19;

Qy 1 MSPNNQNEYILDASSSTVSVDNSVRYVPLANDQTTLLQNMVYKDYLRMSGEENPHLFGNP 60
Db 1 MTSRKENEIINALSIPANSNHAQMWLSTD-----ARIEDSLCTAEG-----NNI 47
Qy 61 ETPISSTVQNGIGVGVQVGLGALGVPPAGQIASPYSFTVGLQMPSTSVSWMKMQVED 120
Db 48 DPFVSASTVQNGINLAGRILGVLPVPPAGQIASPYSFTVGLQMPGR--DPMWIFLEHVEQ 106
Qy 121 LIQKITSVRKLTALAGLOGDGLDVYQKSLKNWLENDRTRARSVVVTVYIALELDFV 180
Db 107 LIROQVTEENTDTALARLQGLGNSFRAYQOQSLLEDLENRRDARTSVLTQYIALELDFL 166
Qy 181 AKIPSAISGOEVLISVYQAANLHLLADASIFGAEWGFTPGCEISFTFYDRQVTRTAQ 240
Db 167 NAMPLPAIRNQEVLVYVYQAANLHLLADASIFGSEGLTSQEIQRVYERQVEKTR 226
Qy 241 YSDYCVKMYNTGLDKLKGNTNAASMLKXHQFREMFLVLDLVALFPNYDTRTYPIETTAQ 300
Db 227 YSDYCARWYNTGLNLRGTNAESWLRYNQFRDLTLGLVLDLVALFPSTYTRVYPMNTSAQ 286
Qy 301 LTRVYTDPIVFNRETSGPCERHSLNSDISFSVESAVIRSHPLFDILSEIEFTTAAQ 360
Db 287 LTRVYTDPIVFNRETSGPCERHSLNSDISFSVESAVIRSHPLFDILSEIEFTTAAQ 345
Qy 361 LPLANTLEYWVGHSHIKYKNTNASSALERNVGTITSNKIKYDILANKDIFQVRSIGADL 420
Db 346 -RWSTQYMYWVGHSHIRSLTSTWTHGNTNTSINPVLTQFTSRDVRTSPAGIN 404
Qy 421 ANYYAQVGVYPAFTLLDKNTGSGVGGFTYSKPHHTMQVCTQNTYNTIDIPPE--NEP 478
Db 405 ILLTPVGVFWARFNN--RNPLNSLRGSLLYTYGT--GVGTQLFDSSETLPPETTRP 460
Qy 479 LSRGYSRLSHITYSFSKKNASSPARYCN---LVPANTHRSADVNTVYSDKTIQIPV 535
Db 461 NYESYHRLSNIRLIS-----GNTLRAPVYSWTHRSADRTNTISSDITQIPLV 509
Qy 536 KAHTLVSGTIVKPGFTGGNLRKTSGLPLAYTSVSVKSPLSQRYRARIYASTNLR 595
Db 510 KSFNLNGSTSVSGPGTGGDIIRTNVNGSVLSGLNPNNTSLQRYRVRYAASQTWVL 569
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:36:05 ; Search time 23 Seconds
(without alignments)
2619.457 Million cell updates/sec

Title: US-10-089-678-1

Perfect score: 6044

Sequence: 1 MSPNNQNYEILDASSTVS.....MSTEGVFNIYESVELVLEE 1167

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4007	66.3	1157	1	US-07-876-280-30
2	4007	66.3	1157	1	US-07-812-180A-2
3	4007	66.3	1157	1	US-08-315-468-2
4	4007	66.3	1157	3	US-07-941-650A-2
5	3662.5	60.6	1149	1	US-07-915-203-2
6	3662.5	60.6	1149	1	US-08-272-887-2
7	3662.5	60.6	1149	2	US-08-789-449-2
8	3152	52.2	1169	1	US-08-315-468-4
9	3123	51.7	1157	2	US-08-532-547-5
10	3123	51.7	1157	2	US-08-379-656B-5
11	3123	51.7	1157	3	US-08-455-838-5
12	3123	51.7	1157	3	US-09-019-809-5
13	3123	51.7	1157	4	US-09-471-177-5
14	3100.5	51.3	1156	4	US-09-002-285-72
15	3100.5	51.3	1156	4	US-09-589-477-72
16	3100.5	51.3	1156	4	US-09-661-322A-28
17	2825.5	46.7	1150	4	US-09-002-285-74
18	2825.5	46.7	1150	4	US-09-589-477-74
19	2715	44.9	1169	1	US-08-542-921-2
20	2715	44.9	1169	2	US-08-880-685-2
21	2715	44.9	1169	2	US-08-880-684-2
22	2710	44.8	1134	4	US-09-002-285-76
23	2710	44.8	1134	4	US-09-589-477-76
24	2549	42.2	1229	1	US-08-100-709-4
25	2549	42.2	1229	1	US-08-176-865-4
26	2549	42.2	1229	1	US-08-474-038-4
27	2549	42.2	1229	2	US-08-779-046-4

28 2549 42.2 1229 2 US-08-881-340-4 Sequence 4, Appli
29 2525 41.8 1227 3 US-09-053-549-2 Sequence 2, Appli
30 2519.5 41.7 1207 1 US-07-951-715A-7 Sequence 7, Appli
31 2519.5 41.7 1207 3 US-08-459-448A-7 Sequence 7, Appli
32 2519.5 41.7 1207 3 US-08-459-595A-7 Sequence 7, Appli
33 2519.5 41.7 1207 3 US-08-459-504B-7 Sequence 7, Appli
34 2519.5 41.7 1207 3 US-08-459-444-7 Sequence 7, Appli
35 2519.5 41.7 1207 3 US-09-053-549-8 Sequence 8, Appli
36 2519.5 41.7 1207 4 US-09-547-422-7 Sequence 7, Appli
37 2459 40.7 1227 1 US-08-448-170-8 Sequence 8, Appli
38 2459 40.7 1227 3 US-08-961-803-9 Sequence 9, Appli
39 2458.5 40.7 1138 1 US-07-973-320-2 Sequence 2, Appli
40 2457 40.7 1227 4 US-09-661-322A-63 Sequence 63, Appl
41 2454.5 40.6 1186 3 US-09-178-252-23 Sequence 23, Appl
42 2454.5 40.6 1186 4 US-09-826-660-23 Sequence 23, Appl
43 2452.5 40.6 1228 4 US-09-661-322A-38 Sequence 38, Appl
44 2441.5 40.4 1138 1 US-07-973-320-4 Sequence 4, Appli
45 2237.5 37.0 1177 3 US-08-754-490-14 Sequence 14, Appl

ALIGNMENTS

RESULT 1

US-07-876-280-30
; Sequence 30, Application US/07876280
; Patent No. 5262158
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M.
; APPLICANT: Cannon, Raymond J.C.
; APPLICANT: Bagley, Angela L.
; TITLE OF INVENTION: No. 5262158el Bacillus thuringiensis Isolates for
; TITLE OF INVENTION: Controlling Acarides
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/876,280
; FILING DATE: 19920430
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M/S 104
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1157 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: kumamotoensis
; INDIVIDUAL ISOLATE: PS50C
; IMMEDIATE SOURCE:
; CLONE: E. coli NM522(pMYC2320) NRRL B-18769
; US-07-876-280-30

Query Match 66.3%; Score 4007; DB 1; Length 1157;
Best Local Similarity 66.1%; Pred. No. 2.3e-285;
Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;

QY 1 MSPNNQNEVEYILDASSSTVSVDNSVRYPLANDQTTLLQNNYKDYLRMSSEGENPELFCNP 60
DB 1 MSPNNQNEVEYILDATPSTVSVDNSVRYPFANEPTDALQNNYKDYLRMSSEGENPELFCNP 60

QY 61 ETFISSSTVOTGIGIVGVLGALGVPFAGQIASFYSFIVGQMFPSSTVSVMEMKQVED 120
DB 61 ETFISSSTVOTGIGIVGVLGALGVPFAGQIASFYSFIVGQMFPSSTVSVMEMKQVED 120

QY 121 LIDOKITDSVKTALAGLQGLDGLDVTYQKSLKNWLENRNDTRARSVVVTVYIALELDPV 180
DB 121 LVDQKIEKYVDKALABELKGLGNALDVTYQSLDWNLENRNDARTSVVSNQFIALDLNFV 180

QY 181 AKIPSPALSGOEVPLISVYAOANLHLLRLDASTFGAENGFTPCIEISTFYDROVTRTAQ 240
DB 181 SSIPSFVSGHEVILLVAYAOANLHLLRLDASTFGAENGFTPCIEISTFYDROVTRTAQ 240

QY 241 YSDVCVMYNTGLDKLKTGNAASMLKYHQFREMTELLVLDLVALFPNYDRTYPIETTAQ 300
DB 241 YSDVCVMYNTGLDKLKTGNAASMLKYHQFREMTELLVLDLVALFPNYDRTYPIETTAQ 300

QY 301 LTREYVTDPIVFNRETSGCFRRWSLNSDISFSEVESAVIRSPHLFDILSIEFYTTTAAQ 360
DB 301 LTRDVTDPDIAFNIVTSGCPNPMSTHSGILFYEVNNVIRPPLFDLILSVEINTSRGG 360

QY 361 LPIANTEYLEVWCHSICKYKNT-NASSALERNYGTITSNKIKYDVLANKDIPQVRSGLAD 419
DB 361 ITLANDAYINTVSGHTLKYRTADSTVTTANYGRITSEK-NSFALEDRODIFENSTVAN 419

QY 420 LANYAQQVYGPYASFTLLDKNTGSGSVGGFTYKPTTMOVCTQNTVIDEIPPE-N-P 478
DB 420 LANYAQQVYGPYASFTLLDKNTGSGSVGGFTYKPTTMOVCTQNTVIDEIPPE-N-P 478

QY 478 LSRGVSHRLSHITSYSPKNSASPARYGNLVPVATHRASVNTVYSDKTIQIPVVAH 538
DB 478 VABSYSHRLSHITSYSPKNSASPARYGNLVPVATHRASVNTVYSDKTIQIPVVAH 538

QY 539 TLVSGTIVIKGFTGGLKARTSSGLAYTSVKSPLSORYARIYASVTLNLRPVT 598
DB 539 TLVSGTIVIKGFTGGLKARTSSGLAYTSVKSPLSORYARIYASVTLNLRPVT 598

QY 599 ISGTRIYSINVKTKMKGDDLTFTFDLATIGTFTFNSYSDSLTVGADSFASGGEVYD 658
DB 599 ISGTRIYSINVKTKMKGDDLTFTFDLATIGTFTFNSYSDSLTVGADSFASGGEVYD 658

QY 659 KPELIPVNAPEABEDLVAKAVNGLTYSKDALQTSVTDYVQVQAANLVECLSDLYP 718
DB 659 KPELIPVNAPEABEDLVAKAVNGLTYSKDALQTSVTDYVQVQAANLVECLSDLYP 718

QY 719 NEKRLMDAVKEAKELVOARNLQDTPGNRINENGWGTGTEIEVAGDVLFPKDRSLRT 778
DB 719 NEKRLMDAVKEAKELVOARNLQDTPGNRINENGWGTGTEIEVAGDVLFPKDRSLRT 778

QY 779 SARREDTETPTLYLQQIDESLLKPYTKLKGFTGSSQDLKEIKLIRANQIVKNVDPN 838
DB 779 SARREDTETPTLYLQQIDESLLKPYTKLKGFTGSSQDLKEIKLIRANQIVKNVDPN 838

QY 839 LLPDVLVNSCGIDRCSEQQVVDANLALENNGENGNMSSDASHPSFHDITGCEIDLNNT 898
DB 839 LLPDVLVNSCGIDRCSEQQVVDANLALENNGENGNMSSDASHPSFHDITGCEIDLNNT 898

QY 899 GFWVFKIPPTNGYATLGNLELVEGSLGSETLERAOQOQOQKMAKGCASAKAYA 958
DB 899 GFWVFKIPPTNGYATLGNLELVEGSLGSETLERAOQOQOQKMAKGCASAKAYA 958

QY 959 AKQALDRLFADYQDQKNSGVMSDMAAQNLVQSTIPVYNDALPEIPGANYTSTFELTN 1018
DB 959 AKQALDRLFADYQDQKNSGVMSDMAAQNLVQSTIPVYNDALPEIPGANYTSTFELTN 1018

QY 949 SKQAVDLRYADYQDQQLNPDVEITDLTAQDLIQISIPVYNNEMFPEIPGMYTKTELTLD 1008
DB 949 SKQAVDLRYADYQDQQLNPDVEITDLTAQDLIQISIPVYNNEMFPEIPGMYTKTELTLD 1008

QY 1019 RLQQAAMLYDLRNLAPNGDFRGLSDWNAISDVNVQQLSDTSVLVPIPNKNSOVSOQFTVQ 1078
DB 1009 RLQQAAMLYDQRNAIPNGDFRGLSDWNAISDVNVQQLSDTSVLVPIPNWDEQVSOQFTVQ 1068

QY 1079 PNRYVLRVTRARKEGVDGYVIRIDGANGQTETLTFNICODDTGVLSADQTSYITKTVFT 1138
DB 1069 PNRYVLRVTRARKEGVDGYVIRIDGANGQTETLTFNICODDTGVLSADQTSYITKTVFT 1128

QY 1139 PSTQVWIDMSSETGVFNIESVBLVLEER 1167
DB 1129 PYTDQWMEISETEGTFTVIESVBLVDVE 1157

RESULT 2
US-07-812-180A-2
; Sequence 2, Application US/07812180A
; Patent No. 5366892
; GENERAL INFORMATION:
; APPLICANT: Foncerrada, Luis R
; APPLICANT: Payne, Jewel M
; APPLICANT: Sick, August J
; TITLE OF INVENTION: No. 5366892el Coleopteran-Active Bacillus
; TITLE OF INVENTION: thuringiensis isolate and a No. 5366892el Gene Encoding a
; TITLE OF INVENTION: Coleopteran-Active Toxin
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Roman Saliwanchik
; STREET: 2421 N.W. 41st Street, Ste A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/07/812.180A
; FILING DATE: 19920102
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, Roman
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1157 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: kumamotoensis
; INDIVIDUAL ISOLATE: PS50C
; IMMEDIATE SOURCE:
; LIBRARY: Lambdagem (TM) - 11 LIBRARY OF LUIS
; LIBRARY: FONCERRADA
; CLONE: 50C
; US-07-812-180A-2

Query Match 66.3%; Score 4007; DB 1; Length 1157;
Best Local Similarity 66.1%; Pred. No. 2.3e-285;
Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;

QY 1 MSPNNQNEVEYILDASSSTVSVDNSVRYPLANDQTTLLQNNYKDYLRMSSEGENPELFCNP 60
DB 1 MSPNNQNEVEYILDATPSTVSVDNSVRYPFANEPTDALQNNYKDYLRMSSEGENPELFCNP 60

QY 61 ETPSSSTVQTGTGIGVGVGALGVPPAGQIASFYFIVGQLWPSSTVSWEIMKQVED 120
DB 61 ETPSSSTVQTGTGIGVGVGALGVPPAGQIASFYFIVGQLWPSSTVSWEIMKQVED 120
QY 121 LIDOKITDTSVRKTAGLQGLGDLGVYQKSLKWLKLNDRNDTRARSVVVTOYIALBLDFV 180
DB 121 LVDQKIEKVYKDFALLKELKGLGNALGVYQSLKWLKLNDRNDTRARSVVVTOYIALBLDFV 180
QY 181 AKIPSPAGSGEVLPLSVYAAQANLHLLRLDASIFGAEMGTPGRISTFYDQVTRTAQ 240
DB 181 SSIPSPAGSGEVLPLSVYAAQANLHLLRLDASIFGAEMGTPGRISTFYDQVTRTAQ 240
QY 241 YSDYCVKNTYGLDKLKTNAASNLKXHQPRREKTLVLVLVALPNNYDTRTPPIETTAQ 300
DB 241 YSDYCVKNTYGLDKLKTNAASNLKXHQPRREKTLVLVLVALPNNYDTRTPPIETTAQ 300
QY 301 LTRVYDTPVNPRTSGGFCRWSLNSDISFSEVESAVIRSPHLDILSEIEFYTTTRAG 360
DB 301 LTRVYDTPVNPRTSGGFCRWSLNSDISFSEVESAVIRSPHLDILSEIEFYTTTRAG 360
QY 361 LPLNTEYLETWGHSIKYKNT-NASSALERNYGTITSNKIKYDILANKDIFQVRLGAD 419
DB 361 LPLNTEYLETWGHSIKYKNT-NASSALERNYGTITSNKIKYDILANKDIFQVRLGAD 419
QY 420 LANYAUVGVPPASFTLLDKNTGSCGVGFTSKPHTTMQVCTQNYNIDEIPPENE-P 478
DB 420 LANYAUVGVPPASFTLLDKNTGSCGVGFTSKPHTTMQVCTQNYNIDEIPPENE-P 478
QY 478 LKSGYSHRLSHITSYFSKNASPARYGNDLPVPAWTHRSADVTNTVYSDKIQTIPVKAH 538
DB 478 LKSGYSHRLSHITSYFSKNASPARYGNDLPVPAWTHRSADVTNTVYSDKIQTIPVKAH 538
QY 539 TLVSGTIVKPGFTGKNTKRTSGPLAVTSVKSPLSQRYRIRYASTNTNLRLFT 598
DB 539 TLVSGTIVKPGFTGKNTKRTSGPLAVTSVKSPLSQRYRIRYASTNTNLRLFT 598
QY 599 ISGTRISYNNVNTMKGDDLLFTNTFDLATIGTAFPSYSDSLTVGADSPASGEVYVD 658
DB 599 ISGTRISYNNVNTMKGDDLLFTNTFDLATIGTAFPSYSDSLTVGADSPASGEVYVD 658
QY 659 KPELIPVNTAFREEDLDVAKKAVNGLPSTSKDALQTSVTDYOVNQAANLVECLSDRLVP 718
DB 659 KPELIPVNTAFREEDLDVAKKAVNGLPSTSKDALQTSVTDYOVNQAANLVECLSDRLVP 718
QY 719 NEKRLMWDVAKELRLVQARNLLQDTGPNRNGENGMTSGTGLEVAEGDVLFDKRLSLRT 778
DB 719 NEKRLMWDVAKELRLVQARNLLQDTGPNRNGENGMTSGTGLEVAEGDVLFDKRLSLRT 778
QY 779 SAREIDTETPTLYQOIDEKLPYTRYKLGPIGSSQDLEIKLIRHRANQIVKQVDPN 838
DB 779 SAREIDTETPTLYQOIDEKLPYTRYKLGPIGSSQDLEIKLIRHRANQIVKQVDPN 838
QY 839 LLPDLVPSNCGGIDRCSEQVVDANLALENNENGENMSDSDSHAPFHDITGTIDLENNT 898
DB 839 LLPDLVPSNCGGIDRCSEQVVDANLALENNENGENMSDSDSHAPFHDITGTIDLENNT 898
QY 899 GIWVGFKIPTNGYATLGNLELVEEGLSGTTLERAQOQEQOQMDKXARRKRGASERYA 958
DB 899 GIWVGFKIPTNGYATLGNLELVEEGLSGTTLERAQOQEQOQMDKXARRKRGASERYA 958
QY 959 AKQAIORLADYQDKLNGVEMSDMLAQNLYQSIPIYVNDALPRIPGNYITSFTELIN 1018
DB 959 AKQAIORLADYQDKLNGVEMSDMLAQNLYQSIPIYVNDALPRIPGNYITSFTELIN 1018
QY 1019 RIQQANLVDLRNAIPNGDFRNLGSDWNATSVDNVQQLSDTSVLVLPNNWSQVSOQFTVQ 1078
DB 1019 RIQQANLVDLRNAIPNGDFRNLGSDWNATSVDNVQQLSDTSVLVLPNNWSQVSOQFTVQ 1078
QY 1079 PNRYVYLRVTAKEGVDGVIIRDGANQTTTLTFCNDDTGVLSDAQDTSYITKTVFT 1138
DB 1079 PNRYVYLRVTAKEGVDGVIIRDGANQTTTLTFCNDDTGVLSDAQDTSYITKTVFT 1138
QY 1139 PSTEQWIDMSETGEGVFNIESVELLEEE 1167

Db 1129 PYTDQWIBISETEGTFFYIESVELVDVE 1157
RESULT 3
US-08-315-468-2
; Sequence 2, Application US/08315468
; Patent No. 5554534
; GENERAL INFORMATION:
; APPLICANT: Michaels, Tracy Ellis
; APPLICANT: Poncerra, Luis
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Process for Controlling Scarab Pests
; TITLE OF INVENTION: with Bacillus thuringiensis isolates
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/315,468
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,941
; FILING DATE: 01 FEB 1993
; APPLICATION NUMBER: 07/828,430
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA: 07/808,316
; APPLICATION NUMBER: 16-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA73.C2
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1157 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: kumamotoensis
; INDIVIDUAL ISOLATE: PS50C
; IMMEDIATE SOURCE:
; LIBRARY: Landdagem (TM) - 11 LIBRARY OF LUIS PONCERRADA
; CLONE: 50C(a)
; US-08-315-468-2

Query Match 66.3%; Score 4007; DB 1; Length 1157;

Best Local Similarity 66.1%; Pred. No. 2.3e-285;

Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;

QY 1 MSPNNQNEYILDASSSTVSNSRYPLANDQTTTLQNNYKYDYLPMSSGEPNLPFNP 60

DB 1 MSPNNQNEYILDASSSTVSNSRYPLANDQTTTLQNNYKYDYLPMSSGEPNLPFNP 60

QY 61 ETPSSSTVQTGTGIGVGVGALGVPPAGQIASFYFIVGQLWPSSTVSWEIMKQVED 120


```
Db 121 LVQKLEKVKDKAELKGLNALDVYQOSLEDLENNDATRSVSNQFALDLPV 180
Qy 181 AKIPSPAIQGVPLLSVVAQAANLHLLLRDASIFGAKWGPTTGEISFYDRQVTRTAQ 240
Db 181 SSIPSPAVSGHEVLLAVVAQAANLHLLLRDASIFGEGWGPTTGEISFYDRQVLTAE 240
Qy 241 YSDYCVKVTYGLDKLKGNAASLKYHOPRREMTLLVLDLVALFPNYDTRTYPIETTAQ 300
Db 241 YSDYCVKVTYGLDKLKGNAASLKYHOPRREMTLLVLDLVALFPNYDTRTYPIETTAQ 300
Qy 301 LTRVYTDPIVFNRETSGGFCRWSLNSDISFSEVESAVIRSPHLPDILSEIEFYFTRAG 360
Db 301 LTRVYTDPIVFNRETSGGFCRWSLNSDISFSEVESAVIRSPHLPDILSEIEFYFTRAG 360
Qy 361 LPLNTEYLEYVWGHISIKYKNT-NASSALERNTGTTISNKIKYDLANKDIPOVRLGAD 419
Db 361 ITLNDAYINWGSHTLKVRRTADSTVYTYANYGRITSEK-NSFALDRDIFRINSTVAN 419
Qy 420 LANYAQUVGPVASTLLDKNTGSGSVGGFTYSKHTTMQVCTQNYNTIDEPENB-P 478
Db 420 LANYAQUVGPVASTLLDKNTGSGSVGGFTYSKHTTMQVCTQNYNTIDEPENB-P 478
Qy 479 LRGYSRHLSHITSYSPKSSNASSPARYGNLPVPAWTHRSADVTNTVYSDKIQIPVVKAH 538
Db 478 VAESYSHRSLHITSYSPKSSNASSPARYGNLPVPAWTHRSADVTNTVYSDKIQIPVVKAH 538
Qy 539 TLVSGTTVTKGPGTGNILKRTSSGPLAVTSVKSPLSQRYRAIRIYASTTNLRLFT 598
Db 536 MLVGGSVVQGPCTGCDILKRTNPSILGTFAVTNGSLSQRYRAIRIYASTTNLRLFT 598
Qy 599 ISGTRIYSINVKTNKGGDDTLTNTDLAGTAPTFSNYSDSLTVGADSPASGGGVYVD 658
Db 595 YLGDTTEKRPNTKMGASLTETTFKFSPIFDQPRETQDKILLSMGDFSSGQEVYID 654
Qy 659 KPELIPVNAATFEABDLDVAKAVNGLFTSKDALOTSVDYQVQNAANLVECLSDLYP 718
Db 655 RIEPIVDETYEADLEAKKAVNALFTWKGLRPGVTDYEVQNAANLVECLSDLYP 714
Qy 719 NEKRLMDVAKEARLVQANLLODTGFRNINENGWGTSGTIEVAGDVLFKDRSLRT 778
Db 715 NEKRLMDVAKEARLVQANLLODTGFRNINENGWGTSGTIEVAGDVLFKDRSLRT 774
Qy 779 SAREIDTETPTLYQOIDEKSLKPYTRYKLPISGSSQDLKLRHRANOIVKSNVDN 838
Db 775 GAREIDTETPTLYQOIDEKSLKPYTRYKLPISGSSQDLKLRHRANOIVKSNVDN 834
Qy 839 LLPDLVPSNCGGIDRCSEQQYDANLALNNGENGNWSSDASHFPHIDTGEIDLMENT 898
Db 835 LLPDLVPSNCGGIDRCSEQQYDANLALNNGENGNWSSDASHFPHIDTGEIDLMENT 888
Qy 899 GIWVFKIPTNGYATLGNLELVEBGLSGETLERAQQOQQOQODKMKARKGASEKAYYA 958
Db 889 GIWVFKIPTNGYATLGNLELVEBGLSGETLERAQQOQQOQODKMKARKGASEKAYYA 948
Qy 959 AKQADRLPADYQDQKLNQVNSDMLAQNLOSIPYVNDALPEIPGANYTSPHETLN 1018
Db 949 SKQAVRLYADYQDQKLNQVNSDMLAQNLOSIPYVNDALPEIPGANYTSPHETLN 1008
Qy 1019 RLQOANLYDLRNPAGDPFRNGLSNWNATSDVNVQQLSDPTSLVATIPNNNSQVSOQFTVQ 1078
Db 1009 RLQOANLYDLRNPAGDPFRNGLSNWNATSDVNVQQLSDPTSLVATIPNNNSQVSOQFTVQ 1068
Qy 1079 PNRYVLRVTRARKEGVGDYVIRIDGANOTETLTFNI CDDDTGVLSADQTSYITKTVFT 1138
Db 1069 PNRYVLRVTRARKEGVGDYVIRIDGANOTETLTFNI CDDDTGVLSADQTSYITKTVFT 1128
Qy 1139 PSTEQWIDMSRTEGVFNIESVELLEER 1167
Db 1129 PYTDQWIBISEGTFTYIESVELLDVE 1157
```

RESULT 5

```
US-07-915-203-2
; Sequence 2, Application US/07915203
; Patent No. 5359048
; GENERAL INFORMATION:
; APPLICANT: Ohba, Michio
; APPLICANT: Iwahana, Hidenori
; APPLICANT: Sato, Reichi
; APPLICANT: Suzuki, No. 5359048ukazu
; APPLICANT: Ogiwara, Katsuooshi
; APPLICANT: Sakanaka, Kazunobu
; APPLICANT: Hori, Hidetaki
; APPLICANT: Asano, Shouji
; APPLICANT: Kawasaki, Tadaaki
; TITLE OF INVENTION: No. 5359048el Microorganism and Insecticide
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/915,203
; FILING DATE: 19920723
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M/K 301
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1149 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-915-203-2
```

```
Query Match 60.6%; Score 3662.5; DB 1; Length 1149;
Best Local Similarity 63.8%; Pred. No. 4.6e-250;
Matches 752; Conservative 109; Mismatches 257; Indels 61; Gaps 21;
```

```
Qy 1 MSPNQNEYIILDASSSTVSNDNSVRYPLANDQTTTLQNNYKYDLPMSGEENBELPCNP 60
Db 1 MSPNQNEYIILDASSSTVSNDNSVRYPLANDQTTTLQNNYKYDLPMSGEENBELPCNP 60
Qy 61 ETPISS-STVOTGIGIVGOVLGAVPPAGQASFYFIVGQLPSPSTSVSWEMIMKQVE 119
Db 61 GTFISADQAVGTGIDIVSTIISGLGIPVLGEVFSILGSLGLLPSNNENWQIPNARVE 120
Qy 120 DLIDQKIDTSVRKLTALAGLQGLGDVYQKSLKNLLENRNDTRARSVVVYQYIALELDF 179
Db 121 ELIDQKIDTSVRKLTALAGLQGLGDVYQKSLKNLLENRNDTRARSVVVYQYIALELDF 180
Qy 180 VAKIPSPAIQGVPLLSVVAQAANLHLLLRDASIFGAKWGPTTGEISFYDRQVTRTA 239
Db 181 RTNMGFSQNTYETPLLPYTAQAASLHLLVNRDQVYIKKENGYPQNDIDLFYKQSVYTA 240
Qy 240 QYSDYCVKVTYGLDKLKGNAASLKYHOPRREMTLLVLDLVALFPNYDTRTYPIETTA 299
Db 241 RYSDHCYQVMTNAGLNKLRGTAKQWVDYNNFRFRNANVAVLDLVALFPNYDARIPLEINA 300
Qy 300 QLTRVYTDPIVFNRETSGGFCRWSLNSDISFSEVESAVIRSPHLPDIL 349
Db 301 ELTRVYTDPIVFNRETSGGFCRWSLNSDISFSEVESAVIRSPHLPDIL 352
```

Qy 350 SEIEFYTTAGLPLANT-EYLEYVGHISIKYKNTNASSALERNYGTITSNKIKYVDLANK 408
Db 353 QEIRMYTS---FRQNGTIEYNYNGGQRLTSLIYIGSSP--NKYSGLVLAGAEDIIIPVQGN 407
Qy 409 DIPVRSIGADLANIYAQVYGPVYASFTLLDKNTGSGVGGFTYKPKHTTMOVCTQYNT 468
Db 408 DIYRV--VMTYIGRYTNSLLGVNPTVP--YFSNNTOK-----TYSKP-----KQFAGGIKT 454
Qy 469 ID---EIPPENEPLSRGYSRHSLSHTSYSPKSNASSPARYGNLVPFAMTHRSADVNTVY 525
Db 455 IDSGBELTYEN---YQSYSHRVSYITSPEIKSTGCTV--LGWVPIFGWTHSSASRNNFY 509
Qy 526 SDKITQIPVKAHTLVSTTVIKGPB-PTGNNILKRTSSGPLAY---TSVSVKPLSORY 581
Db 510 ATKISQIPINKASRTSGCAVNVFQGLYNGGVPVKLSGSGQVNLVRVATDAKG-ASQRY 568
Qy 582 RARIYASTNNLRPLVTS-----GRIYISVNNKTMNKGGDILTENTPLATIGTA-P 633
Db 569 RIRIYASDRAGKT--TISRSBPENPATISAIYNTMTNASTUTSTFAYASGPNL 626
Qy 634 TFSNYSLSLTVGADSPAGSGVYVDKPELIPVNAITFEAEEDLDVAKKAVNGLFTSKDAL 693
Db 627 GISGSRFTDISITKEGAANLYIDRIEFIPVNTLFRAEEDLDVAKKAVNGLFTNEKDAL 686
Qy 694 QTSVTDVQVNOAANLVECLSDLPNEKRLMDVAKEAKRLVQARNLLQDTGPNRINGEN 753
Db 687 QTSVTDVQVNOAANLVECLSDLPNEKRLMDVAKEAKRLVQARNLLQDTGPNRINGEN 746
Qy 754 GWTGSGTGLEVAEGDVLKORSRLTSAREIDTETPTVLYOOIDESLAKPYTRYKLGFI 813
Db 747 GWTGSGTGLEVEGDLVKORSRLTSAREIDTETPTVLYOOIDESLAKPYTRYKLGFI 806
Qy 814 GSSQDLLEIKLRANQIVKXVNPDLPLDVLVNSCGGIDRCSEQQYVDANLALNNGEN 873
Db 807 GSSQDLLEIKLRANQIVKXVNPDLPLDVLVNSCGGIDRCSEQQYVDANLALNNGEN 866
Qy 874 GNMSSDSHAFHIDTGBIDINENTGIIWVVKIPTTNGYATLGNLVELVEGFLSGETLER 933
Db 867 GNMSSDSHAFHIDTGBIDINENTGIIWVVKIPTTNGYATLGNLVELVEGFLSGETLER 926
Qy 934 AQOQEQWQDMARKKASAKYAAKQAIIDRLPADYODQKNSGVMSDMLAAQNLVQS 993
Db 927 AQOQEQWQDMARKKASAKYAAKQAIIDRLPADYODQKNSGVMSDMLAAQNLVQS 986
Qy 994 IPYVNDALPEIPGNYTSTFELNRLQAAWNLXDLRNAIPNGDFRNLGSDWATSDVNV 1053
Db 987 IPYVNDALPEIPGNYTSTFELNRLQAAWNLXDLRNAIPNGDFRNLGSDWATSDVNV 1046
Qy 1054 QOLSSTSLVLPNNNSQVSOQFTVPNTRYVLRVTRARKEGVDGTVIIRDGANOTETLTF 1113
Db 1047 QOLSSTSLVLPNNNSQVSOQFTVPNTRYVLRVTRARKEGVDGTVIIRDGANOTETLTF 1106
Qy 1114 NICDDDTGVLSDQTSYITKTVEPTSPTEQWIDMSETTE 1152
Db 1107 NICDDDTGVLSDQTSYITKTVEPTSPTEQWIDMSETTE 1145

RESULT 6

US-08-272-987-2
; Sequence 2, Application US/08272887
; Patent No. 574750
; GENERAL INFORMATION:
; APPLICANT: Ohba, Michio
; APPLICANT: Iwahana, Hidenori
; APPLICANT: Sato, Reichi
; APPLICANT: Suzuki, No. 5747450ukazu
; APPLICANT: Ogiwara, Katsutoshi
; APPLICANT: Sakanaoka, Kazunobu
; APPLICANT: Hori, Hidetaki
; APPLICANT: Asano, Shouji
; APPLICANT: Kawasugi, Tadaaki
; TITLE OF INVENTION: No. 5747450el Microorganism and Insecticide

NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/272.887
FILING DATE: 08-JUL-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/915,203
FILING DATE: 23-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: M/K 301
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1149 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-272-987-2

Query Match 60.6%; Score 3662.5; DB 1; Length 1149;

Best Local Similarity 63.8%; Pred. No. 4.6e-260; Indels 61; Gaps 21;
Matches 752; Conservative 109; Mismatches 257;
Qy 1 MSPNNQNEYIELDASSSTVSNSVRYPLANDQTTTLQNNYKDYLRMSGBENPELFGNP 60
Db 1 MSPNNQNEYIELDASSSTVSNSVRYPLANDQTTTLQNNYKDYLRMSGBENPELFGNP 60
Qy 61 ETPFSS-STVGTGIGVQVLCALGVPPAGQIASFYSPVIGQLWESSITVVKEMLMKQVE 119
Db 61 GTFISAQDAVGTGIDIVSTIIISGLGIPVLGVSILGLLWPSNNENVMQIFPNRVE 120
Qy 120 DLIDQKITDSVRKTALAGLQGLDGLDLYVQKSLKNLENRNDTRARSVVVTVYIALLEDF 179
Db 121 ELIDQKITDSVRSAIADLANSRVAYEQNALEDWRKNPHSTRSAALVKERPGNAEAIL 180
Qy 180 VAKIPSPAIQOEVPLLSVYAAANLHLLLRDASIFGAEWGFTPGIBISTFYDQVTRTA 239
Db 181 RTNMGSPSQTYETPFLPTTAAQASLHLLVRDQVIGYKGEWGPQNDIDLFTYKEQVSYTA 240
Qy 240 QYSDYCVKMYNTGLDKLTNAASWLKHQFRRENTLLVALPNNYDTTYPLETTA 299
Db 241 RYSDHCQVYNAGLNKLGFTGAKQWVYNNRFRRENNVMVLDLVALPNNYDARIYPLETNA 300
Qy 300 QLTREYVTDPIVFNRETSGGFCRRWSLN---SDI-----SFSEVESAVIRSPHLFDIL 349
Db 301 ELTREIFTDPV-----GSVYTCGSSSTLSIWDIMIPAAALPSFTLEN-LLRKDPFTLL 352
Qy 350 SEIEFYTTAGLPLNNT-EYLEYVGHISIKYKNTNASSALERNYGTITSNKIKYVDLANK 408
Db 353 QEIRMYTS---FRQNGTIEYNYNGGQRLTSLIYIGSSP--NKYSGLVLAGAEDIIIPVQGN 407
Qy 409 DIPVRSIGADLANIYAQVYGPVYASFTLLDKNTGSGVGGFTYKPKHTTMOVCTQYNT 468
Db 408 DIYRV--VMTYIGRYTNSLLGVNPTVP--YFSNNTOK-----TYSKP-----KQFAGGIKT 454
Qy 469 ID---EIPPENEPLSRGYSRHSLSHTSYSPKSNASSPARYGNLVPFAMTHRSADVNTVY 525

Db 455 IDSGEELTYEN---YQSYSHRVSVITSPFEIKSTGGTV--LGVVPFGFTHSSASRNNPIY 509
Qy 526 SDKITQIPVKAHTLVSGTTVIKPG-FTGNNILKRTSSGPLAY--TSVSVKSPLSQRY 581
Db 510 ATKISQIPINKASRTSGAVNMFQGLYNGGVMKLSGSGQVINLRVATDAKG-ASQRY 568
Qy 582 RARIRYASTTNLRFLFVTIS-----GTRIYSINVNKNKGGDLTFTPTPLATIGTA-P 633
Db 569 RIRIRYASDRAGK- -TSSRSPENPATYSASIAFTWMTNSTNASLTYSTFAYSGGPINL 626
Qy 634 TFSNYSDLTVGADSPASGGVYVDKFLI PVNATFEAREDDVDVAKAVNGLPTSKDAL 693
Db 627 GISGSRITFISITKEGAANLIDRIEFPNTLFEAREDDVDVAKAVNGLPTNEKDAL 686
Qy 694 QTSVTDYQVQAANLVECLSDLPYNEKRLMDVAKKRLVQARNLQDTGPNRINGEN 753
Db 687 QTSVTDYQVQAANLIECLSDLPYNEKRLMDVAKKRLVQARNLQDTGFNRINGEN 746
Qy 754 GWTGSTGIEVAGDVLFKDRLSLTSAREIDTETPTLYQOIHDESLKPYTRYKLGPI 813
Db 747 GWTGSTGIEVVEGDVLFKDRLSLTSAREIDTETPTLYQOIHDESLKPYTRYKLGPI 806
Qy 814 GSSODLEIKLRHRANQIVKRVNPNLLPDLVPVNSCGGIDRCSEQOYVDANLLENNGEN 873
Db 807 GSSODLEIKLRHRANQIVKRVNPNLLPDLVPVNSCGGIDRCSEQOYVDANLLENNGEN 866
Qy 874 GMSDSHAFSFHDTGIEDLNENTGIWVFKIPTNGIYATLGNLELVEGFLSGETLER 933
Db 867 GMSDSHAFSFHDTGIEDLNENTGIWVFKIPTNGIYATLGNLELVEGFLSGETLER 926
Qy 934 AQOQEQWQDMARKRGASEKAYAAQAIORLFPADYQDQKLGNGVMSDMLAAQNLVQS 993
Db 927 AQOQEQWQDMARKRAASEKAYAAQAIORLFPADYQDQKLGNGVMSDMLAAQNLVQS 986
Qy 994 IPYVYNDALEIPGMYNTSPFELTNRLQQAANLVDLRNAIPNGDFRNLSDMNATSDVNV 1053
Db 987 IPYVYNDALEIPGMYNTSPFELTNRLQQAANLVDLRNAIPNGDFRNLSDMNATSDVNV 1046
Qy 1054 QQLSDTSVLVLPNNSQVSQFTQPNRYVLRVLTARKEGVDGVIIRDGANQETILTF 1113
Db 1047 QQLSDTSVLVLPNNSQVSQFTQPNRYVLRVLTARKEGVDGVIIRDGANQETILTF 1106
Qy 1114 NICDDRGVLSADQTSYITKTVEPTSPTEQWIDMSET 1152
Db 1107 NICDDRGVLSADQTSYITKTVEPTSPTEQWIDMSET 1145

RESULT 7

US-08-789-449-2
; Sequence 2, Application US/08789449
; Patent No. 5824878
; GENERAL INFORMATION:
; APPLICANT: Obba, Michio
; APPLICANT: Iwahana, Hidenori
; APPLICANT: Sato, Relichi
; APPLICANT: Suzuki, No. 5824878ukazu
; APPLICANT: Ogiwara, Katsutoshi
; APPLICANT: Sakanaka, Kazunobu
; APPLICANT: Hori, Hidetaki
; APPLICANT: Asano, Shouji
; APPLICANT: Kawaasugi, Tadaaki
; TITLE OF INVENTION: No. 5824878el Microorganism and Insecticide
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/789,449

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/272,887

FILING DATE: 08-JUL-1994

APPLICATION NUMBER: US 07/915,203

FILING DATE: 23-JUL-1992

ATTORNEY/AGENT INFORMATION:

NAME: Saliwanchik, David R.

REGISTRATION NUMBER: 31,794

REFERENCE/DOCKET NUMBER: M/K 301

TELECOMMUNICATION INFORMATION:

TELEPHONE: 904-375-8100

TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1149 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-789-449-2

Query Match 60.6%; Score 3662.5; DB 2; Length 1149;

Best Local Similarity 63.8%; Pred. No. 4.6e-260;

Matches 752; Conservative 109; Mismatches 257; Indels 61; Gaps 21;

Qy 1 MSPNNQNEYILDASSSTVSVDNSVRYPLANDQTTTLQNNMYKDYLRMSGENPFLPGNP 60
Db 1 MSPNNQNEYILDASSSTVSVDNSVRYPLANDQTTTLQNNMYKDYLRMSGENPFLPGNP 60
Qy 61 EFTISS-STVGTGIGVGOVLGALGVPPAGQIASFYPIVGOQLPSPSTSVVWEMIMQVYE 119
Db 61 GTFISAQDAVGTGIDIVSTIISGLGIPVLGEVPSILGLATFNENNVQIIPANRVE 120
Qy 120 DLIDQKITDSVRKTALAGLQGLDGLVYQSKLKNLNRNDTRARSVVVQYIALELDF 179
Db 121 ELIDQKILDSVRSRAIADLANSRIAVEYVQNALEDNRKNPHSTRSAALVKERFGNAEIL 180
Qy 180 VAKIPSPALSGQEVPLLSVYQAANLHLLLRDASIFGAEWGFTFCEISTEYDROVTRTA 239
Db 181 RTNMGSPSQTYETPLLEFYAQAASLHLLVWRDVOIYKEMGYPONDIDLFKEQVSYTA 240
Qy 240 QYSDYCVKWNVTGLDKLGTNAASWLKVHOPRENTLLVLDLVALFPNYDTRTYIETTA 299
Db 241 RYSDHCWYNAGLNKLGTCAGKQWVDNRFERNVWVLDLVALFPNYDARIYPLETNA 300
Qy 300 QLTREVYTDPIVFNRETSGGFCRRWSLN-----SDI-----SFSEVESAVIRSHULFDIL 349
Db 301 ELTREIFTDPV-----GSYVTGQSTLISWYDMIPALPSPSTLEN-LLRKPDFFTLL 352
Qy 350 SEIEFTYTRAGLPLANT-EYLEVWVGHISIKYKNTWASSALERNYGTITSNKIKYYDLANK 408
Db 353 QSIEMVTS---FRQNGTIEYNYNGGQRLTGSYIYGSSP--NKYSGVLAGAEDIIIPVQGN 407
Qy 409 DIPVRSILGADLANYYAOVYGVYASFTLLDNRNTSGSVGGFTYKSPHPTTMVCTQNYNT 468
Db 408 DIYRV-VWTVIGRTNSLLGVNPVTF-YFSNNTOK-----TYSKP-----KQFAGGIKT 454
Qy 469 ID---EIPPEPLSRGYSRLSHITTSYFSKQNASPPARYGNLPVPAWTHRSADVNTVY 525
Db 455 IDSGEELTYEN---YQSYSHRVSVITSPFEIKSTGGTV--LGVVPFGFTHSSASRNNPIY 509
Qy 526 SDKITQIPVKAHTLVSGTTVIKPG-FTGNNILKRTSSGPLAY--TSVSVKSPLSQRY 581
Db 510 ATKISQIPINKASRTSGAVNMFQGLYNGGVMKLSGSGQVINLRVATDAKG-ASQRY 568
Qy 582 RARIRYASTTNLRFLFVTIS-----GTRIYSINVNKNKGGDLTFTPTPLATIGTA-P 633
Db 569 RIRIRYASDRAGK- -TSSRSPENPATYSASIAFTWMTNSTNASLTYSTFAYSGGPINL 626

QY 634 TFSNYSLSLTVGADSFASGEVYVVKFELIPVNATFEAEEDLDVAKKAVNGLFTSKKDAL 693
DB 627 GIGSSRTFDISITKEAGANLYIDRIEFIPVNTLFEAEEDLDVAKKAVNGLFTSKKDAL 686
QY 694 QTSVTDYOVQOANLVECLSDLYPNKRLMDVAKKRLVQARNLLQDTGPNRINEN 753
DB 687 QTSVTDYOVQOANLVECLSDLYPNKRLMDVAKKRLVQARNLLQDTGPNRINEN 746
QY 754 GWTGSGTIEVAGDVLKDRSLRLTSAREIDTETPTLYQQIDSLKPYRYKLKGF 813
DB 747 GWTGSGTIEVAGDVLKDRSLRLTSAREIDTETPTLYQQIDSLKPYRYKLKGF 806
QY 814 GSSQDLEIKLIRHRANQIVKVPNDLDPVNSCGGIDRCSEQQYVDANLALNENGEN 873
DB 807 GSSQDLEIKLIRHRANQIVKVPNDLDPVNSCGGIDRCSEQQYVDANLALNENGEN 866
QY 874 GMSDSDSHAFSHDITGIDNENGTIWIPIPTNGVATIGNLELVEBGLSGETLER 933
DB 867 GMSDSDSHAFSHDITGIDNENGTIWIPIPTNGVATIGNLELVEBGLSGETLER 926
QY 934 AQOQEQWQDQWARKRGASAKYAAKQAI DELFADYQDQKLNKSGVEMSDMLAAQNLVQS 993
DB 927 AQOQEQWQDQWARKRGASAKYAAKQAI DELFADYQDQKLNKSGVEMSDMLAAQNLVQS 986
QY 994 IPYVNDALPEIPGNNTSFTLNLRLQOANLVDLRNAINPNDPRNGLSQMNATSDVNV 1053
DB 987 IPYVNDALPEIPGNNTSFTLNLRLQOANLVDLRNAINPNDPRNGLSQMNATSDVNV 1046
QY 1054 QOLSDTSVLVTPNNSQVSOQFTVPYRYVLRVTRKEGVGDGVVIRIDGANOTETLTP 1113
DB 1047 QOLSDTSVLVTPNNSQVSOQFTVPYRYVLRVTRKEGVGDGVVIRIDGANOTETLTP 1106
QY 1114 NICDDTGVLSADQTSYTKTVEFPSTEQWIDMSET 1152
DB 1107 NICDDTGVLSADQTSYTKTVEFPSTEQWIDMSET 1145

RESULT 8
US-08-315-468-4
; Sequence 4, Application US/08315468
; Patent No. 5554534
; GENERAL INFORMATION:
; APPLICANT: Michaels, Tracy Ellis
; APPLICANT: Foncecra, Luis
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Process for Controlling Scarab Pests
; TITLE OF INVENTION: with Bacillus thuringiensis Isolates
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL USA
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/315.468
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/014.941
; FILING DATE: 01 FEB 1993
; APPLICATION NUMBER: 07/828.430
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/808.316
; FILING DATE: 16-DEC-1991

ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA73.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1169 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: kumamotoensis
; INDIVIDUAL ISOLATE: 50C
; IMMEDIATE SOURCE:
; LIBRARY: LambdaGEM-11 library of L. Foncecra
; CLONE: 50C(b)
US-08-315-468-4

Query Match 52.2%; Score 3152; DB 1; Length 1169;
Best Local Similarity 54.0%; Pred. No. 1.4e-222;
Matches 653; Conservative 179; Mismatches 294; Indels 84; Gaps 25;

QY 1 MSPNNQNEYILDDASSSTVSVDNSVRYPLANDOTTTLQNMNMYKDYLRMSSEGENPELPGNP 60
DB 1 MSPNNQNEYILDDASSSTVSVDNSVRYPLANDOTTTLQNMNMYKDYLRMSSEGENPELPGNP 60
QY 61 ETPIS-SSTVQIGIVGQVLGALGVPFAGQIASFYISFVIGQMLPSSSTVSVMIMKQVE 119
DB 61 EVFLSEQDAVKAADIVGKLLTGLGVPFVGPVSVISLYTQIDILMPSKQKSQWEIFMEQVE 120
QY 120 DLIDOKITDSVRKTKALAGLOGLDGLDYQSKLKNWLENDRTRARSVVVQYIALELDF 179
DB 121 ELINOKIAEYARNKALSELEGNYYQLYLTALBEWKNPNRSALDRVNRFEILDSLF 180
QY 180 VAKIPSPAISGOEVELLSVYQAANLHLLLDASIFGAEMGFTPGEISTFYDROVTRTA 239
DB 181 TOYMESFRVTFEVEFLVYVYTAANLHLLLDASIFGEWGLSTSTNNYNRQMKLTA 240
QY 240 QYSDYCVKNTYNGDKLGTNAASLYKHQPEREMTLVLDLVALFPNYDTPTPIETTA 299
DB 241 EYSDHCVMQYETGLAKGSSAKQWIDYNQFEREMTLVLDVVALFSDNTDRTYPLATTA 300
QY 300 QLTREVTDPVFNRETSGGFCRRWLSNSDISFSVESAVIRSPHLPDLSIEFVYTTTA 359
DB 301 QLTREVTDPVFNRETSGGFCRRWLSNSDISFSVESAVIRSPHLPDLSIEFVYTTTA 356
QY 360 GLPLNTEYLVYVGHISIKYKNTNASSALERNYGT-----ITSNKIKYVDLANKDIFQVR 414
DB 357 S--FTSDRYMYWAGHQISYKHIGTSSTFTQYGNQNLQSTN-----FDFNTYDIYKTL 410
QY 415 SIGADLAN-----YYAQYGVPIYASFTLLDKNKGSGVGGFTYKPHHTMQVCTQNTYI 469
DB 411 SNGAVLLDIVPGYTYTFPGMPETEFPMVQNLNTRKT--LTY-KP-ASKDIIDRTDSE 466
QY 470 DEIPPE--NEPLSRGYSRLSHITSYSPSKNASSPARGNLVPWATHRSADVNTVYSD 527
DB 467 LELPPTSGQPNYESYHRLGHIT--FYSSSTST-----YVPVFSWTHRSADVNTVYSG 520
QY 528 KITQIPVVKATLVSGTIVKPGTGNILKRTSSGPLAYTSVSVKSPLSQRYARIRY 587
DB 521 EITQIPGKSSSTIGRNTIIRGRTYGGDLVALTR--IGSCFQMPFESQRIIRY 578
QY 588 AS--TTNRLRF-VTISGTRIYSINVKTK-MNKGD-DLTFTNPDLATIGTFTFSYSDSL 642
DB 579 ASNETSYISLYGLNQSGT-----LKPQYTSNKNENDLYND-----PKYIEPRVI 625

QY 945 MARKGASAKYAAQAIDELFADYODOKLNGVEMSDMLAQNLOSIPVYNDALPE 1004
Db 935 LGRGRABIDRVYLAQAQINHLFPVDYODQQLNPEIGLAENASNLVESISGVYSDTLQ 994
QY 1005 IFCMYNTSFTELNRLOQANWLYDLRNLAPNGDFRNGLSDNWATSQVNVQQLSDTSVLVI 1064
Db 995 IFCINVEIYTELSDRLQOASLYLTSRNAVQNGDFNSCLSDSNWTTMDASVQDCNMEFLVL 1054
QY 1065 PNWNSQVSOQPTVQPHRYVLRVYARKEGVGDGYVIRIDGANTETILTNNICDDDTGVLS 1124
Db 1055 SHWDAQVSOQLRVNPNCKYVLRVYARKEGVGDGYVIRIDGANTETILTNNICDDDTGVLS 1124
QY 1125 ADQTSYITKTVPRTPTFEOWIDMSTEGVFNESVELVLEER 1167
Db 1115 VNDNSYITTEVVPYPTKEMWVEVSESGSYIDSIEPIETQE 1157

RESULT 11

US-08-455-838-5
; Sequence 5, Application US/08455838
; Patent No. 6028246
; GENERAL INFORMATION:
; APPLICANT: Lambert Bart
; APPLICANT: Jansens, Stefan
; APPLICANT: Van Audenhove, Katrien
; APPLICANT: Peferoen, Marix
; TITLE OF INVENTION: NEW BACILLUS THURINGIENSIS STRAINS AND
; TITLE OF INVENTION: THEIR INSECTICIDAL PROTEINS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,838
; FILING DATE: 31-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/01820
; FILING DATE: 12-JULY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 93400949.9
; FILING DATE: 09-APR-1993
; APPLICATION NUMBER: EP 92402358.8
; FILING DATE: 27-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Svensson, Leonard R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 2121-106P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELETYPE: 248345
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1157 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-455-838-5

Query Match 51.7%; Score 3123; DB 3; Length 1157;
Best Local Similarity 52.3%; Pred. No. 1.9e-220;
Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;

QY 1 MSPNNQNEYBILDASSSTSVSDNSVRYPFLANDOTTTLLQNNMYKYDLARMSGENPELFCNP 60
Db 1 MNRNNQNEYBILDAPHCPCPSDDDDRYPLASDPNAAALQNNMYKYDLQWTDYDTSYNP 60
QY 61 ETPISS-STVQTGIGIVGOVLGALGVPPAGQIASFYSPFIVGOLMSPSTSVSWEMIMKQVE 119
Db 61 SLISIGRDAVQVATLVVGRILGALGVPPSGQIVSFYQFLNLTLPVNDTATWEAFMRQVE 120
QY 120 DLIIDQKITDSVRKTALAGLQGLGVYQKSLKWLLENNDTRARSVVVYTOYALELDF 179
Db 121 ELYNQOITFEARNOALARLQGLGDSFNVOYQSLQNLWLRDNDTRNLSSVRAQFIALDLDF 180
QY 180 VAKIPSPALSGQVPLLSVYAQAANLHLLLRDASIFCAEWGPTPGHISTFYDQVTRTA 239
Db 181 VNAIPFPAVNGQVPLLSVYAQAANLHLLLRDASIFCAEWGPTPGHISTFYDQVTRTA 240
QY 240 QYSDYCVKWTNTGDKLKGNTAAWKLKHOFRREMTLLDLVALFPNYOTRTPIETTA 299
Db 241 KYTNYCETWNTGDLRLGNTESMLRYHQFRREMTLLDLVALFPNYOTRTPIETTA 300
QY 300 QLTREVTYDPIVFNRETSGGFCRRNSLNSDISFSEVESAVIRSPHLFDILSEIETTYTRA 359
Db 301 QLTREVTYDPIVFNRETSGGFCRRNSLNSDISFSEVESAVIRSPHLFDILSEIETTYTRA 359
QY 360 GLPLANTEXLYVWVGHSHIKYKNTNASSALERNYGTITSNKIKY--YDLANKDIFVRSAL 416
Db 360 -FPV-SNFMFYWSGHTLRRSYLNSDAVQSDSLGTLITTRATINPGVDGTR-----IBST 413
QY 417 GADLANVYAQVGVPPVASFLL--LDKNTGSGVGGFTYKPHHTMQVCTQNYNTIDEIPP 474
Db 414 AVDFRSALIGIYGVNRASFPVPGGLFNGTTPSPANGG-----CRDLYDNDDELPP 461
QY 475 ENEPLRGYS-HRLSHITSYSPSKN-ASSPARVGNLPVPAWTHRSADVTNTVYSDKTKOI 532
Db 462 DE---STGSSHERLSHVTFPSPQTNQAGSIANAGSVPTVYVTRDVLNNITFNRIQOL 518
QY 533 FVVKATLVSGTIVIKGPGTGNILKRTSGLPAYTSVSKSPLSQRYARIRIYASTTN 592
Db 519 PLVKASAPVSGTIVIKGPGTGNILKRTSGLPAYTSVSKSPLSQRYARIRIYASTTN 578
QY 593 LRLFTVITSGTRIYSINVNKTMKGGDDLTFTNP---DLATIG---TATFPGNSYSLTVGA 646
Db 579 FSIRVLRGVSGIGDVLRLGTMNGOELTYESFPTRFTTTPGPNPPTFTQAOEILTVNA 638
QY 647 DSPASGEVTVVDFPHLPVWATFEAEEDLDVAKAVNGLTSTKDALQTSVTDVQVQAA 706
Db 639 EGVSTGGEYVYIDRIYVVPVNPAREAEEDLEAKKAVASLFTTRTDGLQVNVTDYQVQAA 698
QY 707 NLVECLSDLELYPNEKRLMDAVKAKRLVQARNLQDTGFNRING--ENGWGTSGTGEVA 764
Db 699 NLVCLSDLELYPNEKRLMDAVKAKRLVQARNLQDTGFNRING--ENGWGTSGTGEVA 758
QY 765 EGDVLFDRSLRLTSAREIDTETPTLYYQOIDESELLKPYTRYKLGFGISSQDLKLI 824
Db 759 EGGPFKGRALQASAR---ENYPTIYQVDAVSLKPYTRYKLGFGISSQDLKLI 814
QY 825 RHRANOIVKXVNDLPLPVLVFNSSCGIDRCSCQVYDANLALNNGNGNMSDSDSHAPS 884
Db 815 HHKVLHVLKXVNDLPLPVLVFNSSCGIDRCSCQVYDANLALNNGNGNMSDSDSHAPS 874
QY 885 FHLDTGCEIDLNENTGIWVVPFKIPTTNGYATLGNLEAVEGSLSGTGLERAOQOQOQDK 944
Db 875 SYINTGDLNASVDQGIWVVLKRVTTDGYATLGNLEAVEGSLSGTGLERAOQOQOQDK 934
QY 945 MARKGASAKYAAQAIDELFADYODOKLNGVEMSDMLAQNLOSIPVYNDALPE 1004
Db 935 LGRGRABIDRVYLAQAQINHLFPVDYODQQLNPEIGLAENASNLVESISGVYSDTLQ 994
QY 1005 IFCMYNTSFTELNRLOQANWLYDLRNLAPNGDFRNGLSDNWATSQVNVQQLSDTSVLVI 1064
Db 995 IFCINVEIYTELSDRLQOASLYLTSRNAVQNGDFNSCLSDSNWTTMDASVQDCNMEFLVL 1054
QY 1065 PNWNSQVSOQPTVQPHRYVLRVYARKEGVGDGYVIRIDGANTETILTNNICDDDTGVLS 1124
Db 1055 SHWDAQVSOQLRVNPNCKYVLRVYARKEGVGDGYVIRIDGANTETILTNNICDDDTGVLS 1124
QY 1125 ADQTSYITKTVPRTPTFEOWIDMSTEGVFNESVELVLEER 1167
Db 1115 VNDNSYITTEVVPYPTKEMWVEVSESGSYIDSIEPIETQE 1157

Db 1055 SHMDAQVQQLEVNPNCKYVLRVARKVGGDGVYTRIDGHAHQETLTFNACDYVNGTY 1114
Qy 1125 ADQSYTKTKVEPTSPTEQVWIDMSETGVNIESVELVLEE 1167
Db 1115 VNDNSYITEEVVFPETKHMVSESGSFYDIBFIETQZ 1157

RESULT 12

US-09-019-809-5
; Sequence 5, Application US/09019809
; Patent No. 6143550
; GENERAL INFORMATION:
; APPLICANT: LAMBERT, BART
; APPLICANT: JANSSENS, STEFAN
; APPLICANT: VAN AUDENHOVE, KATRIEN
; APPLICANT: PEPEROEN, MARINX
; APPLICANT: VAN RIE, JEROEN
; APPLICANT: VAN AARSEN, ROEL
; TITLE OF INVENTION: NEW BACILLUS THURINGIENSIS STRAINS AND THEIR
; TITLE OF INVENTION: INSECTICIDAL PROTEINS.
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/019,809
; FILING DATE: 02-FEB-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, LEONARD R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 2121-135P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1157 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-019-809-5

Query Match 51.7%; Score 3123; DB 3; Length 1157;
Best Local Similarity 52.3%; Pred. No. 1.9e-220;
Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;
Qy 1 MSPNNQNEYITLDASSSTVSVDNSVRYPLANDQITTLQNNYKDYLRMSGEGNPELFCNP 60
Db 1 MRRNNQNEYITLDAPHCCPSDDDDVRYPLASDPNNALQNNYKDYLRMTDEDYDTSYNP 60
Qy 61 ETFISS-STVQTGIGVQVIGALGVPPAGQIASPYSFIVGQLMPSSTVSVMIMKQVE 119
Db 61 SLSISGRDAVQVLTAVGRIIGALGVPPSGQIVSFYQPLNTLMPVNDTAWAEMRQVE 120
Qy 120 DLIQKITSVRKTAAGLOGDGLGVQVQSLKRWLENRNDTRARSVVVYQYALELDF 179
Db 121 ELVNNQITEFARNQALRLOGIGDSFNYYQSLQNLADRNDTRNLVSRYAQFIALDIDF 180
Qy 180 VAKIPSPAISQGVPLLSVYAAQANLHLLLRDASIFGABWGFTPGESTPYDQVTRTA 239
Db 181 VNAIPLFAVNGQQVPLLSVYAAQANLHLLLRDASIFGABWGFTPGESTPYDQVTRTA 240

Qy 240 QYSDYCVKXNNTGLDKLKGNAASNLKXHQPRREMTLLVLDLVALFPNYDTRTYPIETA 299
Db 241 KYTNYCETWNTGLRLRGINTESWLRYHQPRREMTLVVLVDVALFPYDYDRLVYTCGSP 300
Qy 300 QLREVVYTDPIVRENETSGGFCREWSANSIDISFSEVSASVIRSHPLDILSIEBPTTRA 359
Db 301 QLREVVYTDPIVFPNPPANVGLCRWGTNPYNTTFSELENAPIRPPLHLPRLNSLTSSNR- 359
Qy 360 GLPLANTLEYLWYVGHISIKYKNTNASSALENYGTYTSNKIKY---YDLANKRDIQVRSIL 416
Db 360 -FPV-SSNPDYWSGHTLRSYLANDSAQVEDSYGLITTRATINPGVDGTR-----IEST 413
Qy 417 GADLANTYQVYGVPIYASFTL--LDKNTGSGVSGFTYSKPHTTMQVCTQYNTIDEIAPP 474
Db 414 AVDFRSALIGYGVNRASFPVGGFLNGTISPGAG-----CRDLYDNDDELPP 461
Qy 475 ENBPLSRGYS-HRLSHITSYSFSKN-ASSPARYGNLVPFAMTHRSADVTNTVYSKIKI 532
Db 462 DE---STGSSSTRHLSHVTFPSQTNQAGSIANAGSVPTVWTRDVLNNTITPNRIQL 518
Qy 533 PVKKAHLVSGTIVIKGPGFTGGNILKRTSSGPLYTSVSKSPLSQRYRARIYASTTN 592
Db 519 PLVKASAPVSGTIVLKGPGFTGGILRRITNGTFTGLRVTVNSPLTQYRLRVRFASG 578
Qy 593 LRLPVTISGTRIYSINVNKNKGDLLTFNTF---DLATIG---TAPFYSNYSDSLTVGA 646
Db 579 PSIRVLRGVSGIDVRLGSTMNRQELTYSEFTREFTTTCGPNPPTFTTQAQELTVNA 638
Qy 647 DSPASGGEVYVDKELIPVNATFEAREDLVAKAVNGLFTSKDALCOTSVTDYOVNOAA 706
Db 639 EGVSTGGYIYDRLEIVFVNPAREAREDLBAKAVASLFTTRDGLQVNVTDYQVQNA 698
Qy 707 NLVCLSDDELAPNEKRMMLDAVKEAKRLVQARNLLQDTGFNRING--ENGWGTSGIYVA 764
Db 699 NLVCLSDDEQYGHDKMLLEAVRAAKRLSRERLLQDPDFNTINSTENGKASNGVTIS 758
Qy 765 EGDVLFKDRSLRLTSARIDTETPTLYYQOIDSLLKPYTRYKLGKFGISSQDLLEIKLI 824
Db 759 EGGPFFKGRALQLASAR----ENPTTYIKVDASVLKPYTRYLDGFPVKSSQDLLEIDL 814
Qy 825 RHRAQIVKVPDNLDPVLPVNSCGGIDRCSEQYVDANLALNNGNGNMSDSSHAPS 884
Db 815 HHKHLVKNVPDNLVSDTYSQSCGSLNRCDEQHVDMQLDAEHHPMDCCAAQTHEPS 874
Qy 885 PHIDTGEIDLNMENTGIWVPKIPITTINGVATLGNLELVEEGLSGSETLERRAQOQOQW 944
Db 875 SYINTGDLNASVDQGIWVLKVRTTGGVATLGNLELVEEGLSGSETLERRAQOQOQW 934
Qy 945 MARIKGASEKAYAAKQAIHRLPADYQOKLNSGVNSDMLAAQNLVQSIPIVYNDALPE 1004
Db 935 LGKRAEIDRVYLAQAQAINELPVDYQDQQLNPEIGLAIINEASNLVESISGVYSDTLQ 994
Qy 1005 IPGKNYTSFTLNTLRLQAQANLYDLRNAIPNGDFRNLSDWNATSDVNVQQLSDTSVLVI 1064
Db 995 IPGKNYTSFTLNTLRLQAQANLYDLRNAIPNGDFRNLSDWNATSDVNVQQLSDTSVLVI 1054
Qy 1065 PNNNSQVSSQQTQVQFNRYRLVTRARKEGVGDGVYIIRDGANQETLTFNICDDDTGVLS 1124
Db 1055 SHMDAQVQQLEVNPNCKYVLRVARKVGGDGVYTRIDGHAHQETLTFNACDYVNGTY 1114
Qy 1125 ADQSYTKTKVEPTSPTEQVWIDMSETGVNIESVELVLEE 1167
Db 1115 VNDNSYITEEVVFPETKHMVSESGSFYDIBFIETQZ 1157

RESULT 13

US-09-471-177-5
; Sequence 5, Application US/09471177
; Patent No. 6448226
; GENERAL INFORMATION:
; APPLICANT: LAMBERT, BART
; APPLICANT: JANSSENS, STEFAN

APPLICANT: VAN AUDENHOVE, KATHRIEN
APPLICANT: PEERHOEN, MARNIX
APPLICANT: VAN RIE, JEROEN
APPLICANT: VAN AARSSEN, ROEL
TITLE OF INVENTION: NEW BACILLUS THURINGIENSIS STRAINS AND THEIR
TITLE OF INVENTION: INSECTICIDAL PROTEINS.
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/471,177
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/019,809
FILING DATE: 02-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, LEONARD R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 2121-135P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1157 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-471-177-5

Query Match 51.7%; Score 3123; DB 4; Length 1157;
Best Local Similarity 52.3%; Pred. No. 1.9e-220;
Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;

QY 1 MSPNNEYEILDASSSTVSNSRYPLANDQTTTLQNNYKDYLEMSGEPELFCNP 60
DB 1 MNPNNNEYEIIDAPHGCPSSDDVRYPLASDPNAAALQNNYKDYLMQTDYDTSYNP 60

QY 61 ETPISS-STVQTGIGVGVGLGALGVPPFAGQIASFYSFVQGLWPSSTVSWMIMKQVE 119
DB 61 SLSISGRDAVOTALTVVGRILGALGVPPFSGQIVSFYQFLNTLPVNDTAIMEAFMEQVE 120

QY 120 DLIDQKITSVRKTAGLGLGDLGVYOKSLKNWLENNDTRARSVVVTVYALELDF 179
DB 121 ELVNOQITEPARNOALRQLGDSFNQYQSLQNLADRNDRNLNSVRAQFALDLDF 180

QY 180 VAKIPSPAISQGVPLISVYQAQANLHLRLDASIFGAEGWGPCEISTFYDQVTRTA 239
DB 181 VNAIPLPFVNGQQVPLLSVYQAQANLHLRLDASIFGAEGWGPCEISTFYDQVTRTA 240

QY 240 QYSDYCVKNTGLDKGTNAASWLYKHOFREMTLLDLVALFPNYDTRYPIETTA 299
DB 241 KTYNYCETWNTGLDRGNTNTSWLYKHOFREMTLLDLVALFPNYDTRYPIETTSNP 300

QY 300 QLTREVTYDPIVFNRETSGGFCRRWSLNSDISFSEVASVIRSHPLDILSRIFETTRA 359
DB 301 QLTREVTYDPIVFNPPANVGLCRWGTNPNTTSELENAFIRPHLPDLRLSLTSSNR- 359

QY 360 GLPLNNTYLEYVWGHISKYKNTNASSALERNYTTISNKIKY---YDLANKDIFQVRS 416
DB 360 -FPV-SSNFMVYNSGHTLRRSYLNSDSAVQEDSYGLTITTRATINPGVDGTR-----TEST 413

QY 417 GADLANYAQQVGVGPVASFLL--LDKNTGSGSVGGFTYKPTHTMQVCTQNTYNTIDISPP 474
DB 414 AVDFRSALIGIYGNRASFPVGGFLFNGTTPSPANGG-----CRDLYDTNDELPP 461

QY 475 ENEPLSRGYS-HLSHLITSYSPSKN-ASSPARVGNLPVPAWTHRSADVNTVYSDKLTQI 532
DB 462 DE---STGSSTHRLSHVTFPSFOTNQAGSIANAGSVFTYVTRDVLNNTITFNRTIQ 518

QY 533 PVVKAHTLVSGTIVKPGFTGNILKRTSSGFLAYTSVSKPLSQRYRARIYASTTN 592
DB 519 PLVKASAPVSGTIVLAGPGFTGGILLRRTTNGFTGLRVTVNSPLTQQYRLRVFASVGN 578

QY 593 LRLFTVTSIGTRIYSINWNTMNGDLDLTNP---DLATIG---TATFSPYSLSLTGVA 646
DB 579 FSIRVLRGVSGICDVRLGSTMNRGQELTYBSPFTRBFTTGPFPNPPFTTQAQILTVNA 638

QY 647 DSPASGGEVYVDFKELI PVNATPEAREEDLVAKAVNGLFTSKDALQTSVTDYQVNOAA 706
DB 639 EGVSTGGEYIIDRIEIVPVNPARAEEDLEAKKAVASLFRTRDGLQVNVTDYQVNOAA 698

QY 707 NLVECLSDELYPNKRMMDAVKAKELVQARNLLQDTGFNRING--ENGWGTSGTGEVA 764
DB 699 NLVSCLSDRQYGEDKMLLEAVRAAKRLSRERNLQDPDFTNTSTEENGKASNGVTIS 758

QY 765 EGDVLPKDRSLRLTSAREIDTETTPYLYQOIDEBSLLKPTRYKLGKFIGSSQDLKLI 824
DB 759 EGGPFKGRALQASAR---ENYPYIYQKVDASVLKPYTRYELDGFVKSSQDLKLI 814

QY 825 RHBAQIVKRVNPDNLDPVLPVNSCGGIDRCSSQVYVDANLALNNGENGMSSDSHAPS 884
DB 815 HHKHLVLRVNPDLNLSVDTYSDGSCGNGINCDHQVDMQDLDAHHHPMDCCERAAQTHERFS 874

QY 885 FHIDTGEIDLNENTGIWVVEFKIPTTNGYATLGNLELVEEGPLSGETTLERAAQOQWQDK 944
DB 875 SYINTGDLASVDQGLVWLKVRTTGYATLGNLELVEVGLSGESLEREQDNKAWAE 934

QY 945 MARKGASERKAYYAKQADRLFPADYQDQKNGSVEMSDMLAQNLYVCSIPYVNDALPE 1004
DB 935 LGRKRAEIDRVYLAQAQAINHLFVDYQDQQLNPEIGLAEINEASNLVESISGVYSDTLQ 994

QY 1005 IPCWNTSFTLNRLOQANLYDLRNAIPNGPFRGLSDWNATSDVNVQQLSDTSVLVI 1064
DB 995 IPIGNTIYETLSDRLOQASLYTSRNAVQNGFNSGLDSWNITMDASVQDQGNHFLVL 1054

QY 1065 PWNNSVSOQFTYQPNRYVLRVTRKGVGDGVIIIRDGANQTETLTFTNCDDDTYGLS 1124
DB 1055 SHWDAQVSQQLRVNFKYVLRVTRKVGSGDGVYTRDGAHQETLTFTNACDYVNGTY 1114

QY 1125 ADQTSYITKTVETPTEQVWIDMSETEGVFNIESVELVEE 1167
DB 1115 VNDNSYITEWVFPETKHMWVEVSESGSYFIDSIEFIETQE 1157

RESULT 14

US-09-002-285-72
Sequence 72, Application US/09002285
Patent No. 6369213
GENERAL INFORMATION:
APPLICANT: Schnepf, H. Ernest
APPLICANT: Wicker, Carol
APPLICANT: Narva, Kenneth E.
APPLICANT: Walz, Michelle
APPLICANT: Stockhoff, Brian
APPLICANT: Muller-Cohn, Judy
TITLE OF INVENTION: Toxins Active Against Pests
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA

Search completed: June 21, 2004, 13:40:18
Job time : 27 secs

Db 455 IDSGEELTYEN---YQSYSHRVSYITSPKISTGTV--LGVVPFGWTHSSASRNPIY 509
Qy 526 SDKITQIPVKAHTLVSGTTWIKGP-FTGNIILKRTSSGPLAY---TSVSVKPLSORY 581
Db 510 ATKISQIPINKASRTSGAVNMFQGLYNGGPMKLSGSGQVINLRVATDAKG-ASQRY 568
Qy 582 RARIRYASTTNLRPLFVTIS-----GTRIYSINVKTNKGGDLTENTPDLATIGTA-P 633
Db 569 RIRIRYASDRAGKF--TISSRSPENPATYSIASIATNTMTNASTNLTSTFYAESGPINL 626
Qy 634 TFSNYSDSLTVGADSFASGGVGVYVDKPELIPVNATFEAEEDLDVAKKAVNGLFTSKDAL 693
Db 627 GISSRTPDISITKEAGANLYIDRIFPVNTLFEAEEDLDVAKKAVNGLFTSKDAL 686
Qy 694 QTSVTDYQVQOANLVECLSDLYPNKRMMDAVKEAKRLVQARNLLQDTGPNRINGEN 753
Db 687 QTSVTDYQVQOANLVECLSDLYPNKRMMDAVKEAKRLVQARNLLQDTGPNRINGEN 746
Qy 754 GWTGSGTGLEVAGDVLFKDRLSLTSAREIDTETPTVLYQOIDSLLKPTRYKLGPI 813
Db 747 GWTGSGTGLEVAGDVLFKDRLSLTSAREIDTETPTVLYQOIDSLLKPTRYKLGPI 806
Qy 814 GSSQDLBIKLRHRANQIVKVPDNLPLDVLVNSCGGIDRCSEQQVYVDANLALNNGEN 873
Db 807 GSSQDLBIKLRHRANQIVKVPDNLPLDVLVNSCGGIDRCSEQQVYVDANLALNNGEN 866
Qy 874 GNMSSDSHAFSHIDTGEIDLNTNGIWNVPKIPPTNGYATLGNLELVEBGLSGETLER 933
Db 867 GNMSSDSHAFSHIDTGEIDLNTNGIWNVPKIPPTNGYATLGNLELVEBGLSGETLER 926
Qy 934 AQOQEQWQDMARKKRGASEKAYAAQADRLPADYQDQKLNKSGVEMSDMLAAQNLVQS 993
Db 927 AQOQEQWQDMARKKRAASEKTYAAQADRLPADYQDQKLNKSGVEMSDMLAAQNLVQS 986
Qy 994 IPYVNDALPEIPGMYNTSTFELNRLQOANWLDLRNAIPNGDFRNLSDWNATSDVNV 1053
Db 987 IPYVNDALPEIPGMYNTSTFELNRLQOANWLDLRNAIPNGDFRNLSDWNATSDVNV 1046
Qy 1054 QQLSDTSVLVIPPNNNSQVQFTVPQNYRYVLRVTRARKEGVDGVVIRIDGANQTETLTP 1113
Db 1047 QQLSDTSVLVIPPNNNSQVQFTVPQNYRYVLRVTRARKEGVDGVVIRIDGANQTETLTP 1106
Qy 1114 NICDDTGVLSADQTSYITKTVEPTPSTEQWIDMSETE 1152
Db 1107 NICDDTGVLSADQTSYITKTVEPTPSTEQWIDMSETE 1145

RESULT 7

US-08-789-449-2
; Sequence 2, Application US/08789449
; Patent No. 5824878
; GENERAL INFORMATION:
; APPLICANT: Ohba, Michio
; APPLICANT: Iwahana, Hidenori
; APPLICANT: Sato, Reichi
; APPLICANT: Suzuki, No. 5824878ukazu
; APPLICANT: Ogiwara, Katsutoshi
; APPLICANT: Sakanaka, Kazunobu
; APPLICANT: Hori, Hidetaki
; APPLICANT: Asano, Shouji
; APPLICANT: Kawasugi, Tadaaki
; TITLE OF INVENTION: No. 5824878el Microorganism and Insecticide
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; City: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/789,449
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/272,887
; FILING DATE: 08-JUL-1994
; APPLICATION NUMBER: US 07/915,203
; FILING DATE: 23-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M/K 301
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1149 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-789-449-2

Query Match 60.6%; Score 3662.5; DB 2; Length 1149;
Best Local Similarity 63.8%; Pred. No. 4.6e-260;
Matches 752; Conservative 109; Mismatches 257; Indels 61; Gaps 21;
Qy 1 MSPNNQNEYETLDASSSTVSVDNSVRYPLANDOTTLLQNMNYKDYLRMSEGENPBLFGNP 60
Db 1 MSPNNQNEYETLDASSSTVSVDNSVRYPLANDOTTLLQNMNYKDYLRMSEGENPBLFGNP 60
Qy 61 ETPISS-STVGTGIGVGVQVLCALGVPPAGQIASFYISVIGOLMPSSTVSVMEMIKQVE 119
Db 61 GTFISAQDAVGTGIDIVSTIISGLGIPVLGEVFSILGSLIGLGMFNNENVMQIFPNRVE 120
Qy 120 DLIDOKITDSVRKTPALAGLQGLDGLDVYKSLKNLNRNDTRARSVVVVOYIALELDF 179
Db 121 ELIDOKILDSVRSRAIADLANSRITAVEYVQNALEDWRKNPHSTRSAALVKERFGNAAIL 180
Qy 180 VAKIPSPAISQOEVPILLSVYAAQANLHLLRDASIFGAEWGFTPGELSTFEYDROVTRTA 239
Db 181 RTNMGSPSQTNVETPLPTFYAQAASLHLLVRDVOIQYKGEWGPQNDIDLFTKBEQSVYTA 240
Qy 240 QYSDYCVKWNVTGLDKLGTNAASWLKYHOFREMTLLVLDLVALFPNYDTTYPIETTA 299
Db 241 RYSDHCQWYNAGLNKLGCTGAKQWVDYRFRKNVAVLDLVALFPNYDARIYPLETNA 300
Qy 300 QLTRFVYTDPIVFNRETSGGFCRRWSLN-----SDI-----SFSVESAVIRSPHLDIL 349
Db 301 ELTREIPTDPV-----GSVYTGSSSTLISWYDMIPALPSPSTLEN--LLRKPDPFTLL 352
Qy 350 SEIEFYTTTRAGLPLANN--EYLEYVWGHISIKYKNTWASSALERNYCTIISNKKIYYDLANK 408
Db 353 QEIRMYTS---FRQNGTTEYNYMGQRLTUSYIYVSSP--NKYSGVLGASDIIPVGON 407
Qy 409 DIFQVRSILGADLANVYAOVGVVPVASFLLLDKQNTSGSGVGGFTYKSPHTTMOVCTQNYNT 468
Db 408 DIYRV--VWTYIGRYTNSLLGVNPTP-YFSNNTOK-----TYSKP-----KQFAGGIKT 454
Qy 469 ID---EIPPENEPLSRGSHRLSHITSYFSKNASSPARYGNLFPVANTHRADVTNTVY 525
Db 455 IDSGEELTYEN---YQSYSHRVSYITSPKISTGTV--LGVVPFGWTHSSASRNPIY 509
Qy 526 SDKITQIPVKAHTLVSGTTWIKGP-FTGNIILKRTSSGPLAY---TSVSVKPLSORY 581
Db 510 ATKISQIPINKASRTSGAVNMFQGLYNGGPMKLSGSGQVINLRVATDAKG-ASQRY 568
Qy 582 RARIRYASTTNLRPLFVTIS-----GTRIYSINVKTNKGGDLTENTPDLATIGTA-P 633
Db 569 RIRIRYASDRAGKF--TISSRSPENPATYSIASIATNTMTNASTNLTSTFYAESGPINL 626

QY 634 TFSVSDSLTVGADSFASGGEVYVDFKELIPVNATPEAREEDLDVAKAVNGLFTSKDAL 693
Db 627 GISSSRTPDISITTEAGAAANLYDIRIBIPVNTLFEAREEDLDVAKAVNGLFTNEKDAL 686
QY 694 QTSVTDYQVQANLVECLSDLYPNEKRWLDVAKVRLVQARNLLODTGPNRNGEN 753
Db 687 QTSVTDYQVQANLVECLSDLYPNEKRWLDVAKVRLVQARNLLODTGPNRNGEN 746
QY 754 GWTGSTGIEVABGDVLFKDRSLRLTSAREIDTETPTLYLQQIDSLKPYRYKLKGP 813
Db 747 GWTGSTGIEVABGDVLFKDRSLRLTSAREIDTETPTLYLQQIDSLKPYRYKLKGP 806
QY 814 GSSQDLLEIKLRHRANQIVKQVNDLPLVLPVNSCGGIDRCSEQQYVVDANLALNNGEN 873
Db 807 GSSQDLLEIKLRHRANQIVKQVNDLPLVLPVNSCGGIDRCSEQQYVVDANLALNNGEN 866
QY 874 GNMSSDSHAFSPHIDTGEIDMLNENTGIWVFKIPTTNGVATLGNLVEEGPLSGETLER 933
Db 867 GNMSSDSHAFSPHIDTGEIDMLNENTGIWVFKIPTTNGVATLGNLVEEGPLSGETLER 926
QY 934 AQOQEQWQDKWARKGASEKAYYAAQAIIDLFLADYQDQKLNKSGVMSDMLAAQNLVQS 993
Db 927 AQOQEQWQDKWARKGASEKAYYAAQAIIDLFLADYQDQKLNKSGVMSDMLAAQNLVQS 986
QY 994 IPYVNDALPEIPGMNYSFTTELTLNLOQAWNLVLRNAILPNGDPRNGLSDMNATSDVNV 1053
Db 987 IPYVNDALPEIPGMNYSFTTELTLNLOQAWNLVLRNAILPNGDPRNGLSDMNATSDVNV 1046
QY 1054 QQLSDTSVLVIPNNSQVSOQPTVQPNRYVLRVTRKRGVGDGVYIIRDGANQTETLTP 1113
Db 1047 QQLSDTSVLVIPNNSQVSOQPTVQPNRYVLRVTRKRGVGDGVYIIRDGANQTETLTP 1106
QY 1114 NICDDDTGVLSDQTSYITKTVFTPTSTEQVWIDMSET 1152
Db 1107 NICDDDTGVLSDQTSYITKTVFTPTSTEQVWIDMSET 1145

RESULT 8
US-08-315-468-4
Sequence 4, Application US/08315468
Patent No. 5554534
GENERAL INFORMATION:
APPLICANT: Michaels, Tracy Ellis
APPLICANT: Poncerra, Luis
APPLICANT: Narva, Kenneth E.
TITLE OF INVENTION: Process for Controlling Scarab Beetles
TITLE OF INVENTION: with Bacillus thuringiensis Isolates
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL USA
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/315,468
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/014,941
FILING DATE: 01 FEB 1993
APPLICATION NUMBER: 07/828,430
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/808,316
FILING DATE: 16-DEC-1991

ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MAY3.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1169 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
STRAIN: kumamotoensis
INDIVIDUAL ISOLATE: 50C
IMMEDIATE SOURCE:
LIBRARY: Lambdagem-11 library of L. Poncerra
CLONE: 50C(b)
US-08-315-468-4

Query Match 52.2%; Score 3152; DB 1; Length 1169;
Best Local Similarity 54.0%; Pred. No. 1.4e-222;
Matches 653; Conservative 179; Mismatches 294; Indels 84; Gaps 25;

QY 1 MSPNNQNEVEIILDASSSTSVSDNSVRYPLANDQTTTLQNNNYKDYLRMSGENPELFGNP 60
Db 1 MSPNNQNEVEIILDASSSTSVSDNSVRYPLANDQTTTLQNNNYKDYLRMSGENPELFGNP 60
QY 61 ETPFIS-SSTVQTGIGVQWLGALGVPPAGQIASFYSPVQGLWPSSTSVSWEMIMKQVE 119
Db 61 EYPLSEQDAVKAADIVGKLLTGLGVPPVGVPIVSLYTQLDILWPSKQKQSEIFMEQVE 120
QY 120 DLIDOKITSVKRTALAGLQGLDGLDVTQKSLKNWLENNDTRASVVTQVIALELDP 179
Db 121 ELINQKIAEYARNKALSEGLGNNGYQLYLTALKEENKPNGRALRDVNRREILDSLF 180
QY 180 VAKIPSPAJSGORVPLLSVYAOANLHLLLRDASIFGAEWGPFGCEISFTFYDROQVTRTA 239
Db 181 TQMPFSFRVTFEVPFLVTTWAANLHLLLRDASIFGAEWGPFGCEISFTFYDROQVTRTA 240
QY 240 QYSDYCVKMYNTGLDKLKGTAASMLKQHFREMTLLVLDLVALFPNYDTRTYPIETTA 299
Db 241 EYSDHCYKMYETGLAKLKGSSAKQWIDYQPREMTLLVLDLVALFPNYDTRTYPIETTA 300
QY 300 QLTREVTDPVFNRETSGCFRRWSLNSDISPSEVESAVIRSPHLFDILSEIFFTTRA 359
Db 301 QLTREVTDPVFNRETSGCFRRWSLNSDISPSEVESAVIRSPHLFDILSEIFFTTRA 356
QY 360 GLPLANTVLEYVWGHSIKYKNTNASSALERNYGT-----ITSNKIKYDLANKDIQVR 414
Db 357 S--FTSDRYPRYWAGHQISYKHIGTSSTFTQMTGNQNLQSTN----PFTNYDVKTL 410
QY 415 SLGADLAN-----YTAQVYGVYPVASFLLDKNTGSGVGGFTYSKPHHTMVOCTQNTY 469
Db 411 SNGAVLLDIVPGYVTPVFPMPETEPFVWVQNLNTRKT--LTY-KP-ASKDIIIDRTDSE 466
QY 470 DRIPPE--NEPLSRGYSHRLSHITSYFSKNASSPARYGNLVPFANTHSADVTNTVSD 527
Db 467 LELPPETSGQPNYSYSHRLGHIT--PIYSSSTST-----YVPVFSWTHRSADLNTVKS 520
QY 528 KITQIPVKAHTLVSGTTVIKGPGTGNILKRTSSGPLAYTSVSKSPLSQRYRARIY 587
Db 521 EITQIPGKSSSTIGRNTYIIKRGYTGGLVALTDR--IGSCFQWIFPESQRIY 578
QY 588 AS--TTNRLP-VTISGTRIYSINVKNT-MNKG-DLTNTNPLDITAGTAFSTNSDSL 642
Db 579 ASNETSYISLYGLNQSQT----LKNQTYSNKNENDLTND-----FKYIEYPRVI 625

Qy 643 TVGADS-----PASGGEVTV-DKELIPVNAFTPEABEDLDVAKAVNGLTSSKDAL 693
Db 626 SYNASSMIQSLSIGITNTNLTIFIPVDTEYAEATDLEAAKAVNALFTNTDGL 685
Qy 694 QTSVTDYQVNAANLVECLSDLELYPNKRLMDAVKAEKELVQARNLLQDTGFNRINGEN 753
Db 686 QPGVTDYEVNAANLVECLSDLELYPNKRLMDAVKAEKELVQARNLLQDTGFNRINGEN 745
Qy 754 GWTGSGTIEVAGDVLKFSRLTSAREIDTETPTLYLQOQDESLLKPYTRYKLKGP 813
Db 746 GWTASTGIEVTEGDAVFKGYRLPGAREIDTETPTLYLQOQDESLLKPYTRYKLKGP 805
Qy 814 GSSQDLLEIKLRBANQIVKRNPDNLDPVLVNSCGGIDRCSEQQVVDANLALENNEN 873
Db 806 GSSQGLIYIRHQTNRIVKRNVPDLDLPDPVPPVNDGRINRCKSEQKYVNSKREVEN--- 862
Qy 874 GNMSSDSHAFSFHIDTGEIDLENVTGIWVFKIPTTNGYATLGNLELVEGSPSGRTLER 933
Db 863 ---SGBAHEFSIPIDTGEIDYENAGIIVGFKITDPGTYATLGNLELVEGSPSGDALER 919
Qy 934 AQOQEQWQDMARKGASKEYAQAQADRLFADYQDQKLNKNGVMSDMLAQNIVQS 993
Db 920 LQKEEQWQKQMTREBETDRRYMASKQAVDRLYADYQDQQLNPVBEITDLTAAQDLIQS 979
Qy 994 IPVYNDALPEICMNYTSPTELTRLQQAQWNLVLDLNAIPNGDPFRNGLSDNATSDVNY 1053
Db 980 IPVYNEMFPEIPGMNTKTELTDRLQQAQWNLVLDLNAIPNGDPFRNGLSDNATSDVNY 1039
Qy 1054 QQLSDTSVLVTPNNNSOVQFTQPNVRYVLRVYRKARKEGVDGVIIRDCANOTETLTF 1113
Db 1040 QQINHTSVLVPWNEQVSKFTQPNVRYVLRVYRKARKEGVDGVIIRDCANOTETLTF 1099
Qy 1114 NICDDTGVLASDQTS---YIT-----KTVEFTPTSTEQWIDMSSETGVFNI 1157
Db 1100 SASDYDTNGMYDQASNTNGYNTNSVYMIKPKRTVDISSVYVQWIRISETEGTFYI 1159
Qy 1158 ESVELVLEE 1167
Db 1160 ESVELVIDVE 1169

RESULT 9
US-08-532-547-5
; Sequence 5, Application US/08532547
; Patent No. 5861543
; GENERAL INFORMATION:
; APPLICANT: LAMBERT, BART
; APPLICANT: JANSSENS, STEPHAN
; APPLICANT: VAN AUDENHOVE, KATRIEN
; APPLICANT: PEPEROEN, MARNIX
; APPLICANT: VAN RIE, JEROEN
; TITLE OF INVENTION: NEW BACILLUS THURINGIENSIS STRAINS AND THEIR
; TITLE OF INVENTION: INSECTICIDAL PROTEINS.
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; FILING DATE: 06-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, LEONARD R.

REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 2121-109P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1157 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-532-547-5

Query Match 51.7%, Score 3123, DB 2; Length 1157;
Best Local Similarity 52.3%, Pred. No. 1.9e-320;
Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;

Qy 1 MSPNNQNEYELDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSEGENPELFGNP 60
Db 1 MNRNNQNEYELIDPHCCCFSDDDVRYPPLASDPNAALQNNYKDYLRMDEDTYSYNP 60
Qy 61 ETFISS-STVGTGIGIVQVLGALGVPAGQIASFYSPFIVGQLMPSSTSVVMEMKQVE 119
Db 61 SLSISGRDAVCTALTVVGRIILGALGVPSPGQIVSFYQFLANTLWPNVDTAIWEAFMRQVE 120
Qy 120 DLIDQKITDSVRKTAGLAGLQGLDGLGVYKSLKNLENRNDTRARSVVVQYIALELDF 179
Db 121 ELVNNQITFEPARNQALRLOGLGDSFNVYQSLQNLADRNDRNLSVVRQAQIALDLDF 180
Qy 180 VAKTIPSAISOEVPILSVVAQAANLHLLLRDASIFCAEKGFTPGELSTPYDROVTRTA 239
Db 181 VNALPLFAVNGQQVPLSVVAQAQVNLHLLLRDASLFGEGNGFTGELSTPYDROVELTA 240
Qy 240 QYSDYCVKYNVTGLDKLGTNAASWKLTKYQFRREMTLLVLDLVALFPNYDTYPIETTA 299
Db 241 KYTNYCETWNTGLDLRGNTSWSLRVYQFRREMTLVLDVVALFPYDVELYPTGSNP 300
Qy 300 QLTREVTYDPIVFNRETSGGFCRAWLSNDISFSEVESAVIRSPHLPDLSIBFTTTRA 359
Db 301 QLTREVTYDPIVFNPPANVGLCRWGTNPYNTFSELENAFIRPPHLPDRNLSTISSNR- 359
Qy 360 GLPLNTEYLEYVWGHSLIKYKNTNASSALERNYGTITSNKIKY---YDLANKD1FOVRSI 416
Db 360 -FPV-SSNFMDFWSGHTLRRSYLNDASVQESYGLITTRATINFGVGTNR----IEST 413
Qy 417 GADLANYAQQVGVYASFTL--LDKNTGSGSVGGFTYSKPHTTMQVCTQNTYNTIDEIPP 474
Db 414 AVDFRSALIGIYGVNRASFVPGGLFNGTTSFANGG-----CRDLVDYDDELFP 461
Qy 475 ENBPLSRGYS-HRLSHITSYSFSKN-ASSPARYGMLPVFAWTHRSADVNTYVSDKITQI 532
Db 462 DE---STGSTRLSHVTFFSFQTNQAGSIANAGSVPTVVMTRRDVLDLNTIPNRITOL 518
Qy 533 PVKANTLVSGTIVIKGPGFTGMLKETSGLAYTSVSVKSPLSORVARIRVASTTN 592
Db 519 PLVKASAPVSGTIVIKGPGFTGGGLRRTTNGTFCGLRVVNSPLTQQRLRVRFSTGN 578
Qy 593 LRLFTVTSITRIYSINVNKTKNGDGLTFNFP---DLATIG---TAFPTSNYSDSLTVGA 646
Db 579 FSIRVLRCGVSIGDVLGSGTWNRGELTYESFPTTGTGFPNFPPTFQAEILTVNA 638
Qy 647 DSPAGSGEVYVDKPELIPVNAFTFAEEDLDVAKAVNGLTSSKDALQTSVTDYQVNOAA 706
Db 639 EGVSTGGEYIDRIEIVPNPAREAEDELEAKAVASLFTTRDGLQVNVTDYQVDOAA 698
Qy 707 NLVCLSDLELYPNKRLMDAVKAEKELVQARNLLQDTGFNRINGEN---ENGWTSCTGIEVA 764
Db 699 NLVCLSDLELYPNKRLMDAVKAEKELVQARNLLQDTGFNRINGEN---ENGWTSCTGIEVA 758
Qy 765 EGDVLPKDRSLRLTSAREIDTETPTLYLQOQDESLLKPYTRYKLKGPFGSSQDLLEIKI 824
Db 759 EGGPFFKGRALQASAR---ENYPTIYQKVDASVLPKPYTRYKLKGPFGSSQDLLEIKI 814

QY 825 RHRANQIVKYNVDNLLPDVLPVNSCGGIDRCSEQQYVDANLALNNGENGMSSDSHAPS 884
DB 815 HHRKVLKYNVDNLLPDVLPVNSCGGIDRCSEQQYVDANLALNNGENGMSSDSHAPS 874
QY 885 PHIDTGEIDNENGTGIVWVFKLPTTNGVATLGNLVEBGPGLSGTTLERAAQOQOQWQDK 944
DB 875 SYINTGDLNASVDQGIWVVKVTRTDTGATLGNLVEBGPGLSGTTLERAAQOQOQWQDK 934
QY 945 MARKGASKEYAAQADRLFADYDQOQKLSGVMSDMLAAQNLVOSIPYVYNDALPE 1004
DB 935 LGRKAEIDRVYLAQAQINHLFVDYDQOQKLSGVMSDMLAAQNLVOSIPYVYNDALPE 994
QY 1005 IPGRYNTSTLTLNRLQANLYDLNAPLNGDFNGLSDNATSDVNVQQLSDTSVLVI 1064
DB 995 IPGINYEYITELSDRLQAASYLYTSRNVAVQNGDFNGLSDNATSDVNVQQLSDTSVLVI 1054
QY 1065 PWNNSQVSOQTPQVQNYRVLTARKVGGVYIIRGDANOTETLTFNICDDDTGVLS 1124
DB 1055 SHWDAQVSOQLRVNPKYVLTARKVGGVYIIRGDANOTETLTFNICDDDTGVLS 1114
QY 1125 ADQTSYITKTVETFTSTEQVWIDMSBTEGVFNIESVLVLEB 1167
DB 1115 VNDNSYITEEVVYFETKHMVVESESGSFYIDGIEFTETQ 1157

RESULT 10

US-08-379-656B-5
; Sequence 5, Application US/08379656B
; Patent No. 5985571
; GENERAL INFORMATION:
; APPLICANT: Lambert Bart
; APPLICANT: Jansens, Stefan
; APPLICANT: Van Audenhove, Katrien
; APPLICANT: Peferoen, Marix
; TITLE OF INVENTION: NEW BACILLUS THURINGIENSIS STRAINS AND
; TITLE OF INVENTION: THEIR INSECTICIDAL PROTEINS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESS: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,656B
; FILING DATE: 23-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/01820
; FILING DATE: 12-JULY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 93400949.9
; FILING DATE: 09-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 92402358.8
; FILING DATE: 27-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Svensson, Leonard R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 2121-104P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEFAX: 248345
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1157 amino acids

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-379-656B-5

Query Match
Best Local Similarity 52.3%; Pred. No. 1.9e-220;
Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;

QY 1 MSPNNQNEIYELDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSEGENPELFGNP 60
DB 1 MNRNQNNEIYELDAPHCPCDDVRYPLASDPNAAQNNYKDYLRMSEGENPELFGNP 60
QY 61 ETFISS-STVQIGIGVQVIGALGVPPAGQIASVYSLVIGOLPSSSTSVVEMINKOYE 119
DB 61 SLSISGRDAVQALTVGRIILGALGVPPSGQISVYQPLCLNTLMPVNDATLAFNAFRQVE 120
QY 120 DLIDQITDSVRKTKALAGLQGLGDVYVQSLKQWLENRNDTRARSVVVTVIALLELP 179
DB 121 ELVNOQITEFARNQALRLQGLGDSFNYYQSLQNLADNRNLSVVAQFIALDLDF 180
QY 180 VAKIPSPALSGOEVPLLSVYAAQANLHLLLDASIFGAEWGFPGEBISTFTDQVTRTA 239
DB 181 VNAIPFVAVNGQVPLLSVYAAQANLHLLLDASIFGAEWGFPGEBISTFTDQVTRTA 240
QY 240 QYSDYCVKMYNTGLDKLTNAASMLKYHOPRENTLVLVLVALPYPNYDTRTYTETTA 299
DB 241 KYTNYCETWYNTGLDRLRGNTESWLRYHOPRENTLVLVLVALPYPNYDTRTYTETTA 300
QY 300 QLTREVVYTDPIVFNRETSGGFCRRWSLNSDISFSEVSAVIRSPHFLDLSEIFYTTRA 359
DB 301 QLTREVVYTDPIVFNRETSGGFCRRWSLNSDISFSEVSAVIRSPHFLDLSEIFYTTRA 359
QY 360 GLPLNTEYLEWVCHSIKYKNTNASSALERNYGTITSNKIKY----YDLANKDIFQVRS 416
DB 360 -FPV--SSNFMQYMSGHTLRRSYLNDASVQEDSYGLITTRATINPGVDCGTNR----IBST 413
QY 417 GADLANYAQQVGVYASFTL--LDKNTGSGVGGFTYSKPHITTMQVCTQNTYNTIDISPP 474
DB 414 AVDFRSALIGYVGNRASFPVGLFNGTSPANGG-----CRDLTDNDELPP 461
QY 475 ENEPLSRGYS-HRLSHITSYSPSKN-ASSPARYGNLVPFAMTHRSADVTNTVYSKITQI 532
DB 462 DE---STGSSHTLSHVTFFSFQTNQAGSIANAGSVPTVVTTRDVLANTITPNTITOL 518
QY 533 PVKKAHTLVSGTIVYKGPFGTGNLTKTSSGLAYTSVSVKSPLSQRTARIRYASTTN 592
DB 519 PLVKASAPVSGTIVYKGPFGTGNLTKTSSGLAYTSVSVKSPLSQRTARIRYASTTN 578
QY 593 LRLPVTISGTIYSINVAKTNKGDLLTFNTP--DLATIG--TAFTFSNYSDSLTVGA 646
DB 579 PSIRVLRGVSGIDVRLGSTMNRGQELYSFTFRETTTGGPFPFPFTTQAGELTVNA 638
QY 647 DSFASGGEVYVDKPELIPVNAITFEAEEDLDVAKKAVNGLFTSKDALQTSVTDYQVNOAA 706
DB 639 EGVSTGGEYIIRIIVPNPAREAEEDLEAAKAVASLFTTRDGLQVNVTDYQVNOAA 698
QY 707 NLVECLSDLEYPNEKRLMADVAKELVQARNLQDTGFNRING--ENGWGTSGTIEVA 764
DB 699 NLVSLCLSDLEYPNEKRLMADVAKELVQARNLQDTGFNRING--ENGWGTSGTIEVA 758
QY 765 EGDVLPKDRSLRSLTSARBIDTETYPYLYQQIDESLLKPYTRYKLGKFGSSQDLRIKLI 824
DB 759 EGGPFFKGRALQASAR-----ENYPTIYQKVDASVLPKPYTRYKLGKFGSSQDLRIKLI 814
QY 825 RHRANQIVKYNVDNLLPDVLPVNSCGGIDRCSEQQYVDANLALNNGENGMSSDSHAPS 884
DB 815 HHRKVLKYNVDNLLPDVLPVNSCGGIDRCSEQQYVDANLALNNGENGMSSDSHAPS 874
QY 885 PHIDTGEIDNENGTGIVWVFKLPTTNGVATLGNLVEBGPGLSGTTLERAAQOQOQWQDK 944
DB 875 SYINTGDLNASVDQGIWVVKVTRTDTGATLGNLVEBGPGLSGTTLERAAQOQOQWQDK 934

Qy	945	MARKRGASFKAYAAKQAIIDRLPADYQDOKLNSGVESDMLAQNLTQSPYTVYNDALPE	1004
Db	935	LGRKRAEIIIDRYLAAKQAIINHLFDYQDQQLNPEIGLABINEASNLVESIGVYSDTLAQ	994
Qy	1005	IPGNATYTSPTELTNRLQAAMNLYDLRNAINPNDPFRGLSDWNATSDVNVQQLSDPTSVLVI	1064
Db	995	IPGNTYIETLSRLQAASVLYTSNRNAVQNGDFNSGLDSWNTTNDASVQDQGNHFLVL	1054
Qy	1065	PNMNSQVSQQPTQPNTRYVLRVTRARKEGVDGGYVIRIDGANQTETLTFFNICDDDTGVL	1124
Db	1055	SHWDAQVSQQLRVNPNCYKYLVRVARKVGGDGYVTVTRDGAHQETLTFFNACDQVNGTY	1114
Qy	1125	ADQTSYTKTKTEFPPTSPTEQVWIDMSSETGVFNIESVELVLLEBE	1167
Db	1115	VNDNSYITKEKVFYPPETKHWVEVSESGSFYIDISIRFIETOE	1157

RESULT 11

```

US-08-455-838-5
; Sequence 5, Application US/08455838
; Patent No. 6028246
; GENERAL INFORMATION:
; APPLICANT: Lambert Bart
; APPLICANT: Jansens, Stefan
; APPLICANT: Van Audenhove, Katrien
; APPLICANT: Peferoen, Marinx
; TITLE OF INVENTION: NEW BACILLUS THURINGIENSIS STRAINS AND
; TITLE OF INVENTION: THEIR INSECTICIDAL PROTEINS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.

```

```

, MIT: 24042
, COMPUTER READABLE FORM:
, MEDIUM TYPE: Floppy disk
, COMPUTER: IBM PC compatible
, OPERATING SYSTEM: PC-DOS/MS-DOS
, SOFTWARE: Patentin Release #1.0, Version #1.25
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/08/455,838
, FILING DATE: 31-MAY-1995
, PRIORITY APPLICATION DATA:
, APPLICATION NUMBER: PCT/EP93/01820
, FILING DATE: 12-JULY-1993
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: EP 93400949.9
, FILING DATE: 09-APR-1993
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: EP 92402358.8
, FILING DATE: 27-AUG-1992
, ATTORNEY/AGENT INFORMATION:
, NAME: Svensson, Leonard R.
, REGISTRATION NUMBER: 30,330
, REFERENCE/DOCKET NUMBER: 2121-106P
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: (703) 205-8000
, TELEFAX: (703) 205-8050
, TELEX: 248345
, INFORMATION FOR SEQ ID NO: 5:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 1157 amino acids
, TYPE: amino acid
, TOPOLOGY: linear
, MOLECULE TYPE: protein
, US-08-455-838-5

```

Query Match	51.7%;	Score 3123;	DB 3;	Length 1157;
Best Local Similarity	52.3%;	Pred. No. 1.9e-220;		
Matches 619;	Conservative 195;	Mismatches 327;	Indels 42;	Gaps 14;

Qy	1	MSPNQNWEYIELDASSSTSVSDNSVRYPLANDQTTTTLONMANYKDYLRMSGEGNPELPGNP	60
Db	1	MNRNQNWEYIELDAPHCGPCSDDDVRYPLASDPNAALQNMANYKDYLOMTDEYDTSYNP	60
Qy	61	ETPFSS-STVOTGIGIVGOVLGALGVPPAGQIASFYSPFVQGLWPSPSTSVSMWIMKQVE	119
Db	61	SLSISGRDAVQATLVVGRILGALGVPPGQIVSPFQFLANTLWPVNDTALWEAPMQVE	120
Qy	120	DLIDQIKTIDSVRKATALAGGLGDGVVQKSLKNWLENRNROTARSVVVTVYIALELOF	179
Db	121	ELVNOQITBEPARNOALARLQGLGDSFNVYQRSLONNLADRNDTRNLSSVVRQAQFIALDLDF	180
Qy	180	VAKIPSPAISGQEWPELLSVYQAANLHLLLRDASIFGAEWGTPGSEISTFYDORVTRTA	239
Db	181	VNAIPLFAVNGOQVPLLSVYQAQVNLHLLLKDAISLFGBCWGPQGEISTFYDORLELTA	240
Qy	240	QYSDYCVKQYNTGLDKLGTNAASMLKYHOFREMTLLVLIDLVALPNYDTRVTRPTETTA	299
Db	241	KYNYCETWTNTGLDRUGRTNTSGWKYHOFREMTLLVLVDVVALFPYDVRILYPGSGNP	300
Qy	300	QLTREVVTDPIVENRETSGFCRRWSLNSDISPSKVESAVIRSPHLLFDILSEIEPYTTRA	359
Db	301	QLTREVVTDPIVENPPANVGLCRKRGNTPNYTPSELENAPIRPPLLFDRLNSLUTISSNR	359
Qy	360	GLPLNNTXYLEYVWGHSHIKYKNTNNAASALERNYTGTTTSNKIKY---YDLANKDIPQVRS	416
Db	360	-FPV-SSNPDYWSGHTLRSLYLNDSAVODSYGLITTRATINPGVDGTRN----IBST	413
Qy	417	GADLANYYAOVGVVPASFTL--LDKNTGSGSVGGFTYSKPHTTWQCTQNYNTIDBIPP	474
Db	414	AVDFRSALIGIYGNRASFPVGGFLNGTTSPPANGG-----CRDLYDNDDELPP	461
Qy	475	ENBPLSRGYS-HRLSHITSYFSKCN-ASSPARTGNLPEFAWTHRSADVNTVYSDKITQI	532
Db	462	DB---STGSSTHLSHVTPSPFQTNQAGSIANAGSVPTVYVTRDVDLNTIIPNRITQL	518
Qy	533	PVKAHTLVSGTIVIKGPGTQGNILKRTSSGPLATYSVSKPPLSQRTARIRYASTTN	592
Db	519	PLVKASAPVSGITVLKGPGTGGGILRRITNGTFTGLRVTVNSPLTQQYRLRVRFPASTGN	578
Qy	593	LRLPVTISGTRIYSINVNKTMNGDDLTFNTP--DLATIG--TAPTSNYSDSLTUGA	646
Db	579	FSRVLKRGVSIQDVLGSTMNGQELUTYESPTRBFTTGPVPPPTFQAQOILTVNA	638
Qy	647	DSPASGEVVVDKFEIIPVNATPEABEDLOAKKAVNGLFTSKKDALQTSVTDYQVNQAA	706
Db	639	EGYSTGEYIIDRIEIVPNPAREABEDLEAKKAVASLPTTRFDGLQVNVTDYQVDOAA	698
Qy	707	NLVECLSDELYPNBKRMMDAVKAEKELVOARNLLQDTGPNRING--ENGWTSSTGLEVA	764
Db	699	NLVCSLDEQYGHDKOMLEAVRAAKLSRERNLLOQDPDNTINSTENGKWKASNGVTIS	758
Qy	765	EGDVLFKDRSLRUTSAREIDTETYPYVLQOIBESLLKPYTRYKLKGPICSSODLEIKLI	824
Db	759	EGGPPFKRALQASAR----ENYPYIYQKVDA5VLKPYTRYRLDGFVKSSODLEIDL	814
Qy	825	RHRANOIVKXVPDNLPLPDVLPVNSCGIDRC5QYVVDANLALENNGENGMSD5SHAPS	884
Db	815	HHKVLHVKNVDPNLVSDTYSDSGSCGINCKDSQHVDQMDAHHPMDCCEAAQTHEFS	874
Qy	885	PHIDTGBIDLNENTGIWVVPKIPITNGYATFLGNLBLEVEGPLSGETLERAQOQOQQHDK	944
Db	875	SYINTGDLNASVDQGIWVLKVRTTDGYATLGNLELVEVGPGLSGESLERQRDNACKWNAE	934
Qy	945	MAKRGKASEKAYAAKQAIIDLRLFADYQDQKLSGVEMKSDMLAAQNLVOSIPYVYNDA	1004
Db	935	LGRKRAEIDRVLAAKQAINHLEFVDYQDQQLNPEIPLGLAEINEASNLVESISGVIY	994
Qy	1005	IPGWNNTSPTELRNLQOAMNLLDLRNAIPNGDFRNLGSDWNATSDVNVQOOLS	1064
Db	995	IPGINTEIYTELSDRLQOASLYLTSRNAVQNGDFNGLSDSWNTTMDASVQDQGNMHLVL	1054
Qy	1065	PNWNSQVSQOQTPQPNRYVYLRVTASKEGVGDGVYIIRDGANQTELTFTNIDDDTGVLS	1124

Db 1055 SHWDAQVSCQRLVNRPNCKVLRVARTKVGCGGYVTRDKGAHQETLTFNACDYDNGTY 1114
Qy 1125 ADQTSYITKTEFTSTEQWIDMSETEGVFNIESVELVLEE 1167
Db 1115 VNDNSYITEEVVPYPTKHMVVESESGSFYIDSIEPIEQE 1157

RESULT 12

US-09-019-809-5
; Sequence 5, Application US/09019809
; Patent No. 6143550
; GENERAL INFORMATION:
; APPLICANT: LAMBERT, BART
; APPLICANT: JANSSENS, STEFAN
; APPLICANT: VAN AUDENHOVE, KATRIEN
; APPLICANT: PEFFEROEN, MARINX
; APPLICANT: VAN RIE, JEROEN
; APPLICANT: VAN AARSEN, ROEL
; TITLE OF INVENTION: NEW BACILLUS THURINGIENSIS STRAINS AND THEIR
; TITLE OF INVENTION: INSECTICIDAL PROTEINS.
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/019,809
; FILING DATE: 02-FEB-1998

CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, LEONARD R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 2121-135P
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345

INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1157 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-09-019-809-5

Query Match 51.7%; Score 3123; DB 3; Length 1157;
Best Local Similarity 52.3%; Pred. No. 1.9e-220;
Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;

Qy 1 MSPNNQVEYILDASSSTSVSDNSVRYPLANDOTTLLQNMNKKDYLRMSGENPELFCNP 60
Db 1 MNRNNQVEYIIDAPHCPCSDDDVRYPLASDPNAAQNMNKKDYLMQWIDEDYDTSYLP 60
Qy 61 ETFISS-STVQIGIGVQVIGALGVPPAGQIASPSYFIVQOLWPSSTSVVEMIMKQVE 119
Db 61 SLSISGRDVAQTALVWGRILGALGVPPSGQIVSPYQFLNLTLPVNDTAIWEAPMRQVE 120
Qy 120 DLIQKITSVRKTAALQGLDGLDYVQSLKWLNRNDYTRARSVVVQYIALELDP 179
Db 121 ELVNNQITTEFARNQALRLQGLGDSFNYYQSLQNLWLRNDYTRNLSVYRAQFIADLDP 180
Qy 180 VAKIPSPAIQGEVPLLSVYAQAANLHLLLRDASIFGAEWGFTPGELSTFFYDROVTRA 239
Db 181 VNAIPLEFVNGQVPLLSVYAQAVNHLHLLLRDASIFGEWGFPGTQGEISTYDROLETA 240

RESULT 13

US-09-471-177-5
; Sequence 5, Application US/09471177
; Patent No. 6448226
; GENERAL INFORMATION:
; APPLICANT: LAMBERT, BART
; APPLICANT: JANSSENS, STEFAN

Qy 240 QYSDYCVQVYNTGLDKLGTNAASWLKHYQFREMTELLVLDLVALPNTDTTRYPIETTA 299
Db 241 KYTNYCETWYNTGLDLRGNTTESWLRHYQFREMTELLVLDLVALPNTDTTRYPIETTSNP 300
Qy 300 QLTREVYTDPIVFNRETSQGPCRRWSLNSDISPESVAVIRSPHLPDILSEIEPYTTA 359
Db 301 QLTREVYTDPIVFNPPANVGLCRRWGTWYNTFSELENAPIRPPHLPDRLSLTISNR- 359
Qy 360 GLPLNNTBYLBYVWGHISIKYKNTNASSALERNYGTITSNKIKY---YDLANKDIFOVRL 416
Db 360 -FPV-SSNFMDFWMSGHTLRRSYLNDASVQEDSYGLITTTTRATINPCVDCGTNR----18ST 413
Qy 417 GADLANYTAQVGVPIASFTL--LDKNTGSGSVGFTYSKPHYTHQVCTQNTYNTIDEIP 474
Db 414 AVDFRSALIGYGVNRASFVPGGLFNGTTPANGG-----CRDLDTNDELPP 461
Qy 475 ENEPISRGYS-HRLSHITSYFSKN-ASSPARYGNLPPAWTHRSADVTNTVYSDKITOI 532
Db 462 DE---STGSSHRLSHVTFPSQTNQAGSIANAGSVPTVWTRRDVLDLNTITPFRITOL 518
Qy 533 FVVKATLLVSGTIVIKGFTGGNLTKRTSGPLAYTSVSVKSPISQRYRARIYASTTN 592
Db 519 FLVKASAPVSGTIVLKGFTGGILRTTNGTPTGLRVTVNSPLTQYRLVRVPASTGN 578
Qy 593 LRLPVTISGTRIYSINWKNKNGDDLTFTNP---DLATIG---TAPTSNYSDSLTVGA 646
Db 579 FSIRVLGSGVIGDVLGSIWNRGQELTYESPFTREFTTTPFPNPFPTTQAEILTVNA 638
Qy 647 DSPASGGVYVDKPELIPVNATPFAEBEDLVAKAVNGLFTSKDALQTSVTDYOVNOAA 706
Db 639 EGVSTGGYIYIDRIEIVPNPAREABEDLEAKAVASLFTTRDGLQVNVTDYQVQAA 698
Qy 707 NLVCLSDLYPNKRMWDAVKEAKRLVQARNLLQDTGFNRING--ENGWGTSGTIGIVA 764
Db 699 NLVCLSDLYPNKRMWDAVKEAKRLVQARNLLQDTGFNRING--ENGWGTSGTIGIVA 758
Qy 765 EGDVLFKDRSLRLTSAREIDTETPTTYLYQQIDESLAKPYTRYKLGKFGSSODLEIKLI 824
Db 759 EGGPFPKGRALQLASAR---ENYPTIYQVYDASVLKPYTRYLDGFPVKSSODLEIDL 814
Qy 825 RHRANOIVQNVDPNLLPDVLPVNSCGGIDRCSEQYVDANTALENNGENKNSSDSHAPS 884
Db 815 HHHKHLVQNVDPNLDVSDTYSGSCGINRCDEQHQVDMQDLAEHHPMDCCAAQTHBFS 874
Qy 885 FHDITGEIDLNENTGIWVFKIPTNGVATLGNLVEEGPLSGETLERAQOOQOQWQDK 944
Db 875 SYINTGDLNASVDQGIWVVKVRTTGDYATLGNLVEEGPLSGETLERAQOOQOQWQAE 934
Qy 945 MARKGASEKAYAAKQAIIDLADYQDOKLNSGVEMSDMLAQNLSQSIPIVYNDALPE 1004
Db 935 LGRKRAEIDRVYLAQAQAINHLFVDYQDQQLNPRIGLAEINERASNLVBSISGVYSDTLQ 994
Qy 1005 IPGMNYSFTLTLNRLOQAWNLXDLRNLAI PNGDFRNLGSDWNATSDVNVQQLSDTSVLVI 1064
Db 995 IPGNIYEIYLSLDRLOQA-SYLYTSRNVAVQNGDFNSGLDSWNTTNDASVQDQGNHFLVL 1054
Qy 1065 PNWNSQVSSQFTVPNTYRVLRTARKEGVGDYVIRDGANGQETTLTFNICDDDTGVLS 1124
Db 1055 SHWDAQVSCQRLVNRPNCKVLRVARTKVGCGGYVTRDKGAHQETLTFNACDYDNGTY 1114
Qy 1125 ADQTSYITKTEFTSTEQWIDMSETEGVFNIESVELVLEE 1167
Db 1115 VNDNSYITEEVVPYPTKHMVVESESGSFYIDSIEPIEQE 1157

APPLICANT: VAN AUDENHOVE, KATRIEN
APPLICANT: PEPEROEN, MARINX
APPLICANT: VAN RIE, JEROEN
APPLICANT: VAN AARSEN, ROEL
TITLE OF INVENTION: NEW BACILLUS THURINGIENSIS STRAINS AND THEIR
TITLE OF INVENTION: INSECTICIDAL PROTEINS.
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA: US/09/471,177
APPLICATION NUMBER: US/09/471,177
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/019,809
FILING DATE: 02-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, LEONARD R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 2121-135P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NOS: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1157 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-471-177-5

Query Match 51.7%; Score 3123; DB 4; Length 1157;
Best Local Similarity 52.3%; Pred. No. 1.9e-220;
Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;

Qy 1 MSPNNQNEYETLDASSSTVSVDNSVRYPLANDQTTTLQNNMYKDYLRMSGEENPELFCNP 60
Db 1 MRRNNQNEYETLDAPHCCPSDDVRYPLASDPNAALQNNMYKDYLTWDBEDYDSYNP 60

Qy 61 ETPISS-STVGTGIGVQVILGALGVPPAGQIASFYFVIGQLWPSSTVSVMIMKQVE 119
Db 61 SLSISGRDAVQTALTWVGRILGALGVPPSGQIVSFYQPLNTLWPNVDTAIWEAFMRQVE 120

Qy 120 DLIQDKITDSVRKTAGLQGLDGLDVYQSLKNWLENRNDTRARSVVVYQYIALELDF 179
Db 121 ELVNNQITTEPARNQALRQGLGSDSNFYVQSLQNLWLRNDTRNLVSVRAQPTALDLDLP 180

Qy 180 VAKIPSPAISQBYPLLSVYAQAANLHLLLRDASIFGAEGWFTPGESTFYDQVTRTA 239
Db 181 VNAIPFVAVQVPLLSVYAQAANLHLLLRDASIFGAEGWFTPGESTFYDQVTRTA 240

Qy 240 QYSDYCVKWNVTGLDKLGTNAASWKLKYHQRREMTLLVLDLVALPNYDTRTYPIETTA 299
Db 241 KYTYNCETWYNTGLDRLRGTNTESWLRVHQFRREMTLLVLDLVALFPYDVRVLTGNSP 300

Qy 300 QLTREVTDTVPFRRETSGGCRWNSLSDISFSEVESAVIRSPHLFDILSEIFPYTTRA 359
Db 301 QLTREVTDTVPFPNPNVGLCRWGTNPYNTFSELENAFIRPPLFDRLNSLTISSNR- 359

Qy 360 GLPLNNTLEYVWCHSKYKNTNASSALERNYGTITSNKIKY---VDLANKDIFQVRS 416
Db 360 -FPV-SSNFMFMDWSGHTLRRSYLNDASVQEDSYGLITTYTRATINPGVDGTR- - - - -TEST 413

Qy 417 GADLANYYAQTGYVPYASFTL--LDXNKGSSVGGFTYSKPHTTMQVCTQWNTYNTIDIPP 474
Db 414 AVDFRSALIGYGVNRASFVPGGLFNGTTSFANGG-----CHOLDYTDNDELPP 461

Qy 475 ENBPLSRGYS-HRLSHITSYSFSKN-ASSPARYGNLPVPAWTHRSADVNTNTVYSKDIQ1 532
Db 462 DE---STGSSTHRLSHVTFPSFTNQAGSIANAGSVPTVYVTRRDVLDLNTTTPNRITQL 518

Qy 533 PVKRAHTLVSGTIVKGGFTGGNLRKRTSGPLAITSVSVKSPULSQRYRARIRVASTTN 592
Db 519 PLVKASAPVSGTIVLKGPGTGGGLRRTTNGTFTGLVTVNTPSLTQOYRLVRFVFASTGN 578

Qy 593 LRLFTTISGTRIYSINVNKTNKGGDLTFNPF---DLATIG---TAPTFNSVDSITVCA 646
Db 579 PSIRVLRGVSGIDVRLGSTMNRGOELTYBSFTRFTTTPGFPNPPFTTQOQELTUNA 638

Qy 647 DSPASGGGVYVDKPELIPVNATFEAREEDLDVAKAVNGLFTSKDALQTSVTDYVNOAA 706
Db 639 EGVSTGGGYIDRIEIVPVNPEAREAEEDLEAKKAVASLPTTRDGLQVNVTDYQVQAA 698

Qy 707 NLVECLSDLEYPNEKRMMDAVKAKRLVQARNLLQDTGPNRNG--ENGWTSGTGIEVA 764
Db 699 NLVSLSDBOYGHDKQMLLEAVRAAKRLSRERNLLQDPDFTNTINSTENGWKAASGVTTIS 758

Qy 765 EGDVLFYDRLSLRTSAREIDTETTYTYLYQOIDSLLKGYTRYKLKGFSSQDLLEIKLI 824
Db 759 EGGPFFKGRALQLASAR----ENPTYIYQKVDASVLKPYTRYRLDGVFKSSQDLLEIKLI 814

Qy 825 RHRANQIVKQVNPMLPDLVPSNCGGIDRCSEQQYVDANLALENNGENNMSSDPAFS 884
Db 815 HHKVKHLVKNVNDLVSDTYSGSCGSLNRCDEHQHVDMDQDAEHHPMDCCAAQTHBFS 874

Qy 885 PHIDTGEIDLNENTGIWVFKIPTTNGVATIGNLELVEBGLSGTSLERAQOQEQWQDK 944
Db 875 SYINTGLNASVDQGIWVLKVRTTDGYATLGNLELVEBGLSGSLEREQRDNKWNAB 934

Qy 945 MARYGASBAYYAAKQNDLADYQDOKLNSGVEMSDMLAONLQVSIPIVYNDALPE 1004
Db 935 LGRKRAEIDRVYLAARQAINHLFVDYQDQQLNPEIGLABINEASNLVSIISGVYSDTLQ 994

Qy 1005 IPGNVYTSFTLTLNRLQOAWNLVYDLRVAIPNGDPRNGLSDMNATSDVNVVQLSDTSVLVI 1064
Db 995 IPGNIYEIYELSDRLQOASVLYTSRAVQNGDFNSGLDSNNTTMDASVQDQGNMFLVL 1054

Qy 1065 PNWNSQVSOQTPVQPNRYVLRVTRARKEGVGDGYVIIRDGANQOTETLTFNICDDDTGVLS 1124
Db 1055 SHWDAQVSQQLRVNPNCKYVLRVTRARKVGGDGYVITRDGAHQHQBTLTFNACDYVNGTY 1114

Qy 1125 ADQTSYITKTVPSTPEQVWIDMSVTEGVNIESVELVLEEE 1167
Db 1115 VNDNSYITEEVVFPYPTKHMVVEVSESEGSFYIDSIFETQBS 1157

RESULT 14
US-09-002-285-72
Sequence 72, Application US/09002285
Patent No. 6369213
GENERAL INFORMATION:
APPLICANT: Schnepf, H. Ernest
APPLICANT: Wicker, Carol
APPLICANT: Narva, Kenneth E.
APPLICANT: Walz, Michelle
APPLICANT: Stockhoff, Brian
APPLICANT: Muller-Cohn, Judy
TITLE OF INVENTION: Toxins Active Against Pests
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA

ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/002,285
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/886,615
FILING DATE: 1-JUL-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/674,002
FILING DATE: 1-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-701C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 1156 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-002-285-72

Query Match 51.3%; Score 3100.5; DB 4; Length 1156;
Best Local Similarity 52.0%; Pred. No. 8.5e-219;
Matches 615; Conservative 196; Mismatches 330; Indels 41; Gaps 13;

QY 1 MSFNQNEVEILDASSTSVNSVRYPIANDQTTILQNNYKDYLMKSEGENPELFGNP 60
DB 1 MNRNNQNEVEIIDAPHCPCSDDDVRYPLASDPNAAALQNNYKDYLMQDDEYDSYNP 60
QY 61 ETPFSS-SVVTQIGIVGOVIGALGVPPAGQASFYSPVIGOLWPSTSVVMEMIKQVE 119
DB 61 SLSISGRDAVQVLTAVGRIALGVPPFGQIVSVFYQFILNTLWPNVDTAIWEAFNRQVE 120
QY 120 DLIDQKITDSVRKALTALQGLDGLDYQKSLKNWLENRNDTRARSVVVTVYIALLEDP 179
DB 121 ELVNNQITFARNQALARLQGLGDSFNVYQRSIQNWLANDRNDTRNLSVVRQAFIALDLDF 180
QY 180 VAKIPSPALSGQVEPILLSVYQAQANLHLLLRDASIFGAEWGFTPGSEISTFTDQVTRTA 239
DB 181 VNAIPFAVNGQVPLLSVYQAQVNLHLLLRDASLFGEGWGFTQGEISTFTYDROLELTA 240
QY 240 QYSDYCVKAVNTGLDKLKGTAASWLKYHQFREMTLLVLVLPNPDYDTRYPYETTA 299
DB 241 KYNYCETWYNTGLDRGNTESWLRYHQFREMTLVLDVVALFPYDVRVLYPTGSPN 300
QY 300 QLTREYVTPDIPVNRSTSGGFCRWSLNSDISPSEVESAVIRSPHLDILSEIEFTYTRA 359
DB 301 QLTREYVTPDIPVNPANVGLCRWGTNPVNTSELENAPIRPPLHFDRLNSLTSSNR- 359
QY 360 GLPLANTEVLEYVWGHISIKYKVTNASSALERNYGTITSNKIKY--YDLANKDIFQVRS 416
DB 360 -FPV-SNFMXYMSGHTLRSLYLNDSAVQSDSYGLITTTTATINPGVDGTR----IEST 413
QY 417 GADLANYAQQVGVVASFLL--LDKNTGSGSGVGGFTYSKPHTTMQVCTQNYNTIDEIPP 474
DB 414 AVDFRSALGIYGVNRRASFVPGGLFNGTTSFANGG-----CRDLYTNDLELPP 461
QY 475 ENPEPLSRGXS-HRLSHITSYSFSKN-ASSPARYGNLPFAWTHRSADVNTVYSDKITQI 532
DB 462 DE---STGSSTHRLSHVTPSPQTNQAGSIANAGSVPTVYVTRRDVDLNNITPNTITQL 518
QY 533 PVVKAHTLVSGTIVIKGPGFTGGNLIKRTSSGGLATYSVSKSPLSQRYRARIYASTTN 592

DB 519 PLVKASAPVSGTIVLKPGFTGGILRRITNGTFTGLRVTVNSPLQRRVRVRFPASSGN 578
QY 593 LALFVTISGTRIYSINVNKNKGGDLTFTNTDLATIGT-----AFTFSNYSDSLTVGAD 647
DB 579 PSIRILRGNTSIAYQRFSGTMRGQBLTVESFVTSFTTNQSDLPFTFTQAOENLILAE 638
QY 648 SPASGGEVYVDFELIPVNATFEABEDLVAKAVNGLFTSKDALQTSVTDYQVNOAAN 707
DB 639 GVTSGSEYFIDRIEIPVNPAREABEDLEAAKAVANLFRTRDGLQVNVTDYQVNOAAN 698
QY 708 LVECLSDLELPNEKEMLDVAKELVQARNLQDQTFGNRING--ENGWTSGTGLVAV 765
DB 699 LVSCLSDEQYGHDKKMLLEAVRAAKRLSRERLLQDPDFTNTINSTENGKANGVTISE 758
QY 766 GDVLFDRSLRLTSAREIDTETPTVLYQOIDEISLKPTRYRYLKGFGIGSSQDLKILIR 825
DB 759 GGPFFKGRALQLASAR---ENYPTVIYQVDASVLKPVTRYRLDGFVKSSQDLIDLIH 814
QY 826 HRANQIVKNVDPNLLPDVLPVNSCGGIDRCSSGOQYVDANLALENGNGNMSDSDSHAFSP 885
DB 815 HHKVELVKNVDPNLLVSDTYSOGSCSINRCDBQHVDQMDAEHPMDCCCAAQTHFSS 874
QY 886 HLDTCRIDLNENTGIWVFKIPTTNGYATLGNLELVEEGPLSGETLERRAQOQOQWODKM 945
DB 875 YINTGDLRASVDQGIWVVKVTTDGYATLGNLELVEGVLGSGESLEREQDRNAKNAEL 934
QY 946 ARRGASAKAYAAQOALDRLPADYODQKXNGSVMESDMLAQNLOSIPYVYNDALPEI 1005
DB 935 GKRRABIDRVYLAQAQINHLFVDYQDQQLNPEIGLAEINBEASNLVESISGVYSDTLLOI 994
QY 1006 PGNTTSPFELNRLQOAWNLVDLNAIPNGDFRNLGSLDNATSDVNVQOLSDTSLVIP 1065
DB 995 PGINTEIYTELSDRLQOASLYTSRNAVQNGDFNSGLDSMNITMDASVQOQDGNHFLVLS 1054
QY 1066 NNSOVSOQFTQPNRYVRLVTRARKEGVGDGVIIIRDGANOTETILTFINCDDDTGVLSA 1125
DB 1055 HMDAQVSQLRVNPNCKYVLRVTRARKVGGDGYVTRDGAHHEITLTFNACDYDNGTYV 1114
QY 1126 DQTSYTKTVEFTPTQVWIDMSETEGVFNIESVELLEBE 1167
DB 1115 NDNSYITEBVFVFPETKMMVSESGSFYIDSIFBIETQE 1156

RESULT 15
US-09-589-477-72
Sequence 72, Application US/09589477
Patent No. 6570005
GENERAL INFORMATION:
APPLICANT: Schnepf, H. Ernest
APPLICANT: Wicker, Carol
APPLICANT: Narva, Kenneth E.
APPLICANT: Walz, Michelle
APPLICANT: Stockhoff, Brian
TITLE OF INVENTION: Toxins Active Against Peats
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/589,477
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: US 08/886,615
; FILING DATE: 1-JUL-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/674,002
; FILING DATE: 1-JUL-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: WA-701C1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (352) 375-8100
; TELEFAX: (352) 372-5800
; INFORMATION FOR SEQ ID NO: 72:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1156 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-589-477-72

Query Match          51.3%; Score 3100.5; DB 4; Length 1156;
Best Local Similarity 52.0%; Pred. No. 8.5e-219;
Matches 615; Conservative 196; Mismatches 330; Indels 41; Gaps 13;

QY      1 MSPNNQNEYEILDASSSTVSVDNSVRYPLANDQTTTLQNNMYKDYLRMSEGENPELPCNP 60
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      1 MNRNNQNEYEIIDAPHCCPSDDVDRVPLASDPNNAQLQNNMYKDYLTQMTDEDYDTSYNP 60
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      61 ETPISS-STVGTGIGVQVLGALGVPPAGQIASFYSGVLQAPSSSTVSVMEMKQVE 119
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      61 SLSISGRDAVQATLTWGRILGALGVPPSGQIVSFYQPLANTLWPNVDTAIWEAPMRQVE 120
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      120 DLIQDKITDSVRKTAALAGLQGLDGVYQKSLKNWLENRNDTRARSVVVTOYTALELDP 179
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      121 ELVNQQTTFEARNQALARLQGLGDSFNVYQSRQLQNLADRNDTNLSVVRAQPIALDLDLP 180
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      180 VAKIPSPAISGOBYPLLSVYAAQANLHLLLRDASIFGAERGFPGGEISTFYDQVTRTA 239
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      181 VNALPLFVANGQQVPLLSVYAAQVNLHLLLRDASLFGEGHGFGTQGEISTFYDQLELTA 240
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      240 QYSDYCVKQYNTGLDKLKGTTNAASWLKYHQPRREMTLLVLDLVALFPNYDTRTYPIETTA 299
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      241 KYTWYCTWYNTGLDRLRGTTVESWLRVYHQPRREMTLVLDLVALFPYVDVRLYPTGSPN 300
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      300 QLTREVYTDPIVFNRETSGGRCRWSLNSDLSFSEVESAVIERSPHLFDILSEIRFYTTTRA 359
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      301 QLTREVYTDPIVFNPPANVGLCRRWGTMTPYNTFSLENAFIRPPLHFDRLNSLTISSNR- 359
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      360 GLPLNNTEYLEYVWCHSFKYKNTNASSALERNYGTITSNKIKY---YDLANKDIPQVRSL 416
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      360 -FPV-SNFMFDYNGHTLRRRLYDLNDSAVQEDSYGLIITRATINPGVDGTR-----TEST 413
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      417 GADLANTYAAQVYGVYPSFTLL--LDKNTGSGSVGGFTYSKPHHTMQVCTQNYNTIDEIIPP 474
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      414 AVDFRSALIGIYGNRASFPVGGLPNGTTPSPANGG-----CRDLYDNDDELPP 461
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      475 ENEPLSRGYS-HRLSHITSYSPSKN-ASSPARYGNLPIVFATTHRSADVNTVYSDKITQI 532
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      462 DE---STGSSYTHRLSHVTFPFSQTNQAGSLANAGSVPTVYWRDRVDVLNNTITFNRIQTQL 518
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      533 PVVKAHTLVSGTWTIKGPGFTGEMILKETSGLAYTSVSKSPLSQRYRARIYASVTN 592
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      519 PLVKAAPVSGTIVLKGPGFTGGGLRLRTTNGTFTGLVTVNSPLTIQIRIVRVFPASSGN 578
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      593 LRLPVTISGTRIYSINNVKTNKNGDILTFNTFDLATICI-----AFTFSNYSDSLTVGAD 647
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      579 PSIRILRGNTSIAYQRFGSTWNRGOELTYESPVTFSEFTTNGSDLPFTTQAGENLITLAE 638
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      648 SFASGGEVYVDKFLIPVNAFTFRAEEDLDVAKAVNGLPTSEKDALQTSVTDYDQVNOBAN 707
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```